University of Alberta

Functional and Topological Analysis of Acyl-CoA:Diacylglycerol Acyltransferase 2 From *Saccharomyces cerevisiae*

by

Qin Liu

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Plant Science

Department of Agricultural, Food and Nutritional Science

Spring 2011
Edmonton, Alberta

ABSTRACT

Acyl-CoA:diacylglycerol acyltransferase (EC 2.3.1.20, DGAT or DAGAT) is a membrane protein found mainly in the endoplasmic reticulum (ER). It catalyzes the final step in the biosynthesis of triacylglyerol (TAG or TG), which is the principal repository of fatty acids for energy utilization and membrane formation. Several lines of evidence have indicated that DGAT has a substantial effect on carbon flux into TAG. DGAT has at least two discrete family members (DGAT1 and DGAT2) with different physiological roles. High-resolution structures of both DGATs, however, are absent due to difficulties in purification. In order to gain insight into structural and functional relationships of DGATs, a functional DGAT2 protein from the yeast *Saccharomyces cerevisiae* (ScDGAT2, also known as Dgalp) was selected. The structural and functional role of cysteine residues in ScDGAT2 was studied using site-directed mutagenesis (SDM) in combination with chemical modification. Although ScDGAT2 is susceptible to thiol-modifying reagents, none of the cysteines are essential for the catalytic activity or involved in structure support though disulfide linkages. Inhibition of DGAT activity by thiol-specific modification was localized to cysteine^{314}, which is in the proximity of a highly conserved motif in DGAT2s. Thus, cysteine^{314} may reside in a crucial position near a possible active site or related to proper protein folding. The functional importance and topological orientation of signature motifs

in ScDGAT2 were also studied using the same methods. Both the N- and C-termini of ScDGAT2 are oriented toward the cytosol. A highly conserved motif, ^{129}YFP^{131}, and a hydrophilic segment exclusive to ScDGAT2, reside in the ER and play essential roles in enzyme catalysis. In addition, the strongly conserved H^{195}, which may be part of the active site of DGAT2, is likely embedded in the membrane. Although ScDGAT2 has a topology similar to that of murine DGAT2, there are striking differences which suggest that the topological organization of DGAT2 is not ubiquitously conserved.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor Dr. Randall Weselake for allowing me to pursue my PhD studies in his group, for giving a stimulating and flexible environment that helps me to develop independent research skills, and for providing me opportunities to discuss my research at conferences. I would like to also acknowledge my supervisory committee members, Drs. Brian Sykes, Nat Kav and Ryan McKay, for their advice though my studies which make me a better scientist. I'm grateful to Dr. Christopher McMaster and Michael Gaenzle for serving at my dissertation committee.

This thesis could not have been finished without the help from many people. A big thank you goes to Crystal Snyder for her tireless support all the way through my program and above all, being a great friend. I thank Dr. Rodrigo Siloto for his academic help in many ways and valuable discussions. I also thank all other past and present lab members of Weselake lab for their help. I appreciate Dr. Zhanyun Guo for his patience answeing my questions about chemical modification through emails. I also appreciate Drs. Sten Stymne and Ulf Stahl for providing the yeast strains. I am grateful for support provided by administrative staff Dr. Urmila Basu, Laura Smith and Jody Forslund of Department of Agricultural, Food and Nutritional Science. Financial assistance from Natural Science and Engineering Research Council of Canada, the Canada Research

Chairs Program and the Canada Foundation for Innovation is also gratefully acknowledged.

Finally, I would like to thank my husband Weifeng. He has been my strength, my very best friend and a pillar of support. I am also grateful to my parents and my brother for their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACAT: acyl-CoA:cholesterol acyltransferase

ARAT: acyl-CoA:retinol acyltransferase

ASO: antisense oligonucleotide

AWAT: acyl-CoA:wax alcohol acyltransferase

BCA: bicinchoninic acid

BSA: bovine serum albumin

cDNA: complementary DNA

CoA: coenzyme A

DAG: *sn*-1,2-diacylglycerol

DGAT or DAGAT: acyl-CoA:diacylglycerol acyltransferase

DMSO: dimethyl sulfoxide

DNA: deoxyribonucleic acid

DTT: dithiothreitol

DTNB: 5, 5'-dithiobis-2-nitrobenzoate

EDTA ethylene diamine tetraacetate

EMS: ethyl methane sulfonate

EST: expressed sequence tag

ER: endoplasmic reticulum

FA: fatty acid

FFA: free fatty acid

G3P: *sn*-glycerol-3-phosphate

GPAT: glycerol-3-phosphate acyltransferase

LPAAT: lysophosphatidic acid acyltransferase

NAFLD: non-alcoholic fatty liver disease

HEPES: 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid

HRP: horseradish peroxidase

IA: iodoacetamide

MAM: mitochondria-associated membranes

MBOAT: membrane-bound O-acyltransferase

2-ME: 2-mercaptoethanol

MGAT: acyl-CoA:monoacylglycerol acyltransferase

NCBI: National Center for Biological Information

NEM: N-ethylmaleimide

ORF: open reading frame

PAP: phosphatidic acid phosphatase

PEG-mal: mPEG5000-maleimide

PVDF: polyvinylidene fluoride

RACE: rapid amplification of cDNA ends

RNA: ribonucleic acid

RT-PCR: reverse transcription-polymerase chain reaction

SCD: stearoyl-CoA desaturase

SDM: site-directed mutagenesis

SDS: sodium dodecyl sulphate

SDS-PAGE: sodium dodecyl sulphate-polyacrylamide gel electrophoresis

TAG or TG: triacylglycerol

TLC: Thin Layer Chromatography

TMD: transmembrane domain

VLDL: very low-density lipoprotein

WS: wax ester synthase

WT: wild type

YNBD: yeast nitrogen base with dextrose

The two-character code preceding each *DGAT* gene and polypeptide indicates the organism of origin as follows:

Aa, *Aedes aegypti*; Ac, *Aspergillus clavatus*; Ah, *Arachis hypogaea*; An, *Aspergillus niger*; Ao, *Aspergillus oryzae*; At, *Arabidoposis thaliana*; Bf, *Branchiostoma floridae*; Bn, *Brassica napus*; Bt, *Bos Taurus*; Ce, *Caenorhabditis elegans*; Ci, *Coccidioides immitis*; Cn, *Cryptococcus neoformans*; Dm, *Drosophila melanogaster* ; Dd, *Dictyostelium discoideum*; Dr, *Danio rerio*; Ea, *Euonymus alatus*; Gm, *Glycine max*; Gz, *Gibberella zeae*; Hs, *Homo sapiens*; Jc, *Jatropha curcas*; Lb, *Laccaria bicolor*; Md, *Monodelphis domestica*; Mm, *Mus musculus*; Mt, *Medicago truncatula*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*; Nf, *Neosartorya ischeri*; Nt, *Nicotiana tabacum*; Nv, *Nematostella vectensis*; Oe, *Olea europaea*; Os, *Oryza sativa*; Pf, *Perilla frutescens*; Pm, *Penicillium marneffei*; Pn, *Phaeosphaeria nodorum*; Pp, *Physcomitrella patens*; Pt, *Populus trichocarpa*; Ps, *Picea sitchensi*; Rc, *Ricinus communis*; Rg, *Rhodotorula glutinis*; Rn, *Rattus norvegicus*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ss, *Sus scrofa*; Ta, *Trichoplax adhaerens*; Tg, *Toxoplasma gondii*; Tm, *Tropaeolum majus*; Ts, *Talaromyces stipitatus*; Um, *Ustilago maydis*; Ur, *Umbelopsis ramanniana*; Vf, *Vernicia fordii*; Vg, *Vernonia galamensis*; Vv, *Vitis vinifera*; XL, *Xenopus laevis*; YL, *Yarrowia lipolytica*; Zm, *Zea mays*.

1. Introduction

Triacylglycerol (TAG or TG) is a major storage lipid used as an energy reserve and a source of fatty acid (FA) building blocks for membrane lipid synthesis. Basic TAG metabolism is conserved from bacteria to humans and is essential for many physiological processes. The classic pathway for synthesis of TAG includes three sequential acyl-transfers from acyl-CoA to a glycerol backbone (Weselake, 2005). Acyl-CoA:diacylglycerol acyltransferase (DGAT or DAGAT, EC 2.3.1.20) catalyzes the terminal acyl-transfer reaction. DGAT activities reside mainly in two membrane bound polypeptides, known as DGAT1 and DGAT2. It is believed that DGAT reactions may regulate the flow of carbon into TAG (Perry et al., 1999; Stals et al., 1994).

Both DGAT1 and DGAT2 are associated mainly with the endoplasmic reticulum (ER), but they have distinct roles in TAG bioassembly in different organisms. Knockout of the *DGAT1* gene in mice (*Dgat1*$^{-/-}$) resulted in a substantial decrease of TAG levels in tissues whereas *DGAT2*-deficient mice (*Dgat2*$^{-/-}$) died prematurely (Smith et al., 2000; Stone et al., 2004). In common domesticated plants, DGAT1 plays a major role in oil production (Li et al., 2010). In plants producing unusual FAs, however, DGAT2 is suggested to be important for selective accumulation of these FAs in seed oil (Kroon et al., 2006; Shockey et al., 2006).

Given the importance of DGATs in TAG biosynthesis in animals and plants, understanding the molecular mechanism of DGAT action is important to better manipulate TAG biosynthesis. Indeed, DGAT is regarded as a key target for treating a variety of diseases triggered by disorders of TAG metabolism (Chen and Farese, 2000; Chen et al., 2002b; Choi et al., 2007; Yu et al., 2005), and the enzyme is an attractive tool for boosting the accumulation of seed TAGs for nutritional or industrial applications.

To date, DGAT action has been partially characterized though studies of gene structure and expression pattern, protein primary structure, biochemical properties, membrane topology and subcellular localization. No high-resolution structural information, however, is currently available for either DGAT1 or DGAT2 due to the difficulties in the purification of these membrane-bound proteins.

The *DGAT1* gene encodes a polypeptide of around 500 amino acid residues in length and contains multiple hydrophobic segments. Compared to DGAT1, DGAT2 is generally less than 400 amino acid residues in length, less hydrophobic and predicted to have a simpler topological organization (Shockey et al., 2006). The overall goal of this PhD project was to obtain new structural and functional information on a recombinant DGAT2 using a combination of

molecular biology and biochemistry approaches, based on the following hypotheses:

1) Structural and functional features of a recombinant DGAT2 with substantial biological activity resemble those of DGAT2 under physiological conditions.
2) Since microsomal DGAT activity has been shown to be sensitive to cysteine-specific inhibition (Kamisaka et al., 1993; Lozeman et al., 2001; Sauro and Strickland, 1990), cysteine residues may have a role in DGAT2 structure/function.
3) Signature motifs in DGAT2 are involved in enzyme function.

Overall, the results presented in this study provide the evidence that *Saccharomyces cerevisiae* DGAT2 is reliable enzyme candidate for structure/function studies. The inhibition of ScDGAT2 activity by thiol-modification was localized to a single cysteine residue at position 314, which is close to a putative active site or a region important for proper protein folding. In addition, several signature motifs were found to be essential for ScDGAT2 catalysis, and their topological orientations were determined. These insights contribute to our understanding of structure and function in DGAT2 at a molecular level, and will benefit the experimental design of further structure/function studies on this class of protein.

2. Literature Review

2.1 TAG assembly and the role of DGAT

Triacylglycerol (TAG or TG), a major class of neutral lipid, is a non-polar acyl triester of glycerol with fatty acids (FAs). The occurrence of TAG is widespread in eukaryotes including animals, fungi and plants as well as prokaryotes. TAG has multiple functions but serves principally as a reservoir of FAs for energy utilization and membrane formation (Karantonis et al., 2009). In animals, besides energy storage in adipocytes, TAG is important for many other essential physiological processes such as intestinal fat absorption, lactation, attenuation of lipotoxicity and a water barrier function of the skin surface (Farese et al., 2000; Yen et al., 2008). Excessive accumulation of TAG, however, is related to a variety of diseases including obesity, type II diabetes, coronary heart disease and non-alcoholic fatty liver disease (NAFLD) (Millar and Billheimer, 2006; Mulhall et al., 2002; Stone et al., 2004). In the case of plants, TAG is the main component of seed oils, providing the energy during seed germination (Graham, 2008). TAG is also reported in fungi and some bacteria as an energy storage compound (Alvarez and Steinbuchel, 2002; Sandager et al., 2002). TAGs produced by plants and microorganisms have the potential to replace non-renewable crude oil in several industrial applications including the production of fuel, lubricants and paints (Lung and Weselake, 2006; Siloto et al., 2009a; Snyder

et al., 2009). Modification of TAG production in plants and microorganisms is becoming a rapidly developing field in biotechnology.

TAG biosynthesis has been extensively studied since the 1950s and may involve several pathways (Dahlqvist et al., 2000; Sorger and Daum, 2003; Yen et al., 2002). The classical pathway leading to TAG formation is known as the Kennedy or *sn*-glycerol-3-phosphate (G3P) pathway (Weiss and Kennedy, 1956), which has been described in different categories of organisms including animals, plants, fungi and bacteria (Alvarez and Steinbuchel, 2002, Athenstaedt and Daum, 2006). This pathway involves three sequential acyl-transfers from acyl-CoA to glycerol backbone (Figure 2.1). It starts with the acyl-CoA-dependent acylation of G3P to form lysphosphatidic acid (LPA) catalyzed by acyl-CoA:*sn*-glycerol-3-phosphate acyltransferase (GPAT). The second acylation, catalyzed by acyl-CoA:lysophosphatidic acid acyltransferase (LPAAT), leads to the formation of phosphatidic acid (PA). Following dephosphorylation to form diacylglycerol (DAG), the third acylation is catalyzed by a acyl-CoA:diacylglycerol acyltransferase (DGAT, EC 2.3.1.20).

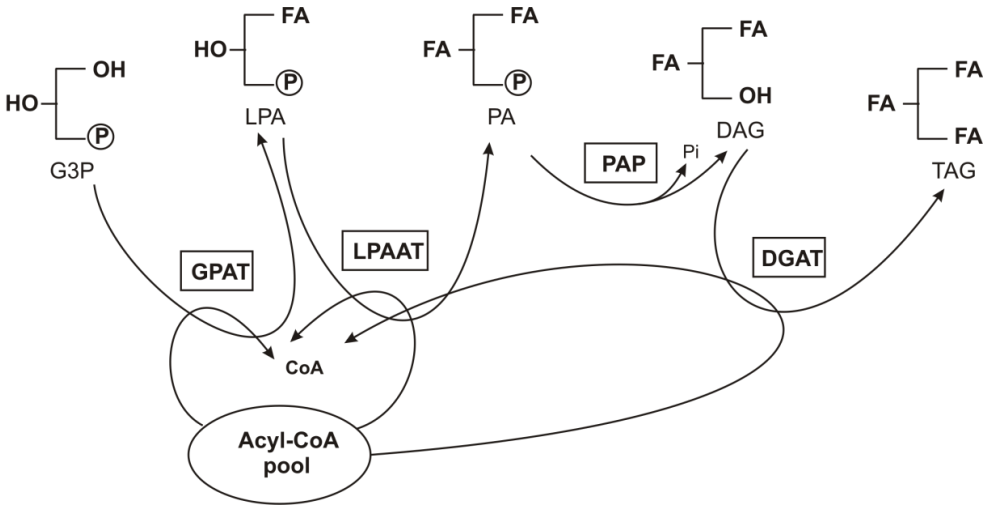


Figure 2.1. Sequential acylation of the glycerol backbone in the acyl-CoA-dependent synthesis of triacylglycerol (TAG) in eukaryotes or prokaryotes (Alvarez and Steinbuchel, 2002; Athenstaedt and Daum, 2006). FA, fatty acid; G3P, sn-glycerol-3-phosphate; GPAT, acyl-CoA:*sn*-glycerol-3-phosphate acyltransferase; LPA, lysophosphatidic acid; LPAAT, acyl-CoA:lysophosphatidic acid acyltransferase; PAP, phosphatidate phosphatase; DAG, sn-1,2-diacylglycerol; DGAT, acyl-CoA:diacylglycerol acyltransferase. Enzymes are enclosed in rectangles.

As the enzyme catalyzing the terminal and committed step in the acyl-CoA-dependent synthesis of TAG, DGAT has been proposed to play a key role in determining the carbon flux into TAG. For instance, Stals et al. (1994) observed that the rate of TAG synthesis in rat hepatocytes is controlled by the affinity of DGAT for acyl-CoA. Similarly, a close correlation between DGAT activity and oil accumulation has been reported in oilseed crops such as canola (*Brassica napus*) (Perry et al., 1999). Therefore, increasing our understanding of the molecular mechanisms involved in DGAT catalysis could lead to a development of novel strategies for manipulation of TAG assembly in different organisms. In humans, DGAT has been considered as a promising target for designing drugs to combat disorders related to TAG overaccumulation as described above. On the other hand, DGAT is also an attractive tool in developing new strategies to produce oils enriched in desirable FAs in microorganisms and oleaginous plant species (Lung and Weselake, 2006, Snyder et al., 2009). Although DGAT activity has been reported since the 1950s (Weiss and Kennedy, 1956), purification of a membrane DGAT to homogeneity has been unsuccessful. Advances on structure and function of DGATs were limited until two microsomal DGATs (DGAT1 and DGAT2) were characterized at the molecular level. This review begins with the discovery of different DGAT isoforms, and then focuses on the recent progress in probing structural and functional aspects of DGAT1 and DGAT2. DGAT

applications in both therapeutics and oilseed engineering, which may be aided by insights into structure/function relationships of DGATs, are also discussed.

2.2 Molecular characterization of DGATs

To date, two known microsomal DGATs, designated DGAT1 and DGAT2, have been identified with orthologues in animals, plants and microorganisms. Several other related proteins have also been characterized including bifunctional enzyme wax ester synthases (WS) exhibiting DGAT activity and soluble DGATs (Kalscheuer and Steinbuchel, 2003; Stoveken et al., 2005).

The first *DGAT* cDNA (known as *DGAT1*) was cloned from mouse in 1998 (Cases et al., 1998) based on sequence homology between an expressed sequence tag (EST) and a previously isolated acyl-CoA:cholesterol acyltransferase (ACAT), a DGAT-related enzyme involved in cholesterol ester biosynthesis (Chang et al., 1993). Functional expression of the mouse cDNA encoding DGAT1 in insect cells proved that this protein possessed DGAT activity (Cases et al., 1998). Orthologues from a number of other species, such as thale cress (*Arabidopsis thaliana*), baker's yeast (*Saccharomyces cerevisiae*) and human (*Homo sapiens*) were quickly identified (Bouvier-Nave et al., 2000b; Oelkers et al., 2002). *DGAT1* was not found in the genome of most of yeast organisms although it has been suggested that a *DGAT1* was found in *Yarrowia*

lipolytica, a nonconventional yeast growing on unusual carbon source such as hydrocarbons (J-M Nicaud et al., personal communication).

In the study to characterize *A. thaliana* mutants, a *DGAT1* mutant (AS11) induced by ethyl methane sulfonate (EMS), was identified at both molecular and biochemical levels (Katavic et al., 1995; Routaboul et al., 1999; Zou et al., 1999). This *A. thaliana* mutant had reduced seed oil content, an altered FA profile and delayed seed maturation because of an insertional mutation in the *DGAT1* gene. DGAT activity could still be detected, however, suggesting the existence of another enzyme was responsible for the remaining enzyme activity. This work was in agreement with a reverse genetic study in mice in which *DGAT1*-deficient mice (*Dgat1*$^{-/-}$) had normal plasma TAG levels and abundant lipids in adipose tissue (Smith et al., 2000). Collectively, these findings indicated the possibility of an alternative mechanism for synthesizing TAG, which was subsequently shown to be mediated by a second class of DGAT referred to as DGAT2.

DGAT2 polypeptide from lipid bodies of the fungus *Umbelopsis ramanniana*, (formerly known as *Mortierella ramanniana*) was first identified through a combination of partial purification and gel electrophoresis of column chromatography fractions enriched in DGAT activity (Lardizabal et al., 2001). Polypeptide excised from the electrophoresis gel was partially sequenced and the information used to produce degenerate primers which were in turn used to isolate

DGAT2 cDNA. In the same study, orthologues from *Caenorhabditis elegans* and *S. cerevisiae* were also cloned and shown to encode proteins with DGAT activities. These DGAT2s shared little or no homology with known *DGAT1* and *ACAT* genes. These results formed the basis for subsequent cloning of this new DGAT family member from mouse and human (Cases et al., 2001). Knockout of *DGAT2* (*dgat2*$^{-/-}$) in mouse led to early death and skin barrier abnormalities (Stone et al., 2004). Shortly thereafter, Sandager et al. (2002) reported that DGAT2 (also known as Dga1p), was the major enzyme contributing to TAG synthesis in *S. cerevisiae*. While *DGAT1* is only characterized in *Y. lipolytica* as described above, *DGAT2* is the only DGAT characterized in most varieties of yeast. Molecular characterization of *DGAT2*s from plants has been of a challenge, probably because early attempts to functionally express a *DGAT2* cDNA from model plant *A. thaliana* were not successful (Burgal et al., 2008; Shockey et al., 2006; Zhang et al., 2009). This may have hindered the interest in further studies on other plant orthologues as will be discussed in Chapter 3. On the other hand, DGAT2 polypeptide has higher sequence divergence compared to DGAT1 which may have also caused the difficulties in cloning other gene orthologues from plants. In 2006, two novel *DGAT2* cDNAs were cloned from castor bean (*R. communis*) and tung tree (*Vernicia fordii*) which produce unusual seed oils rich in ricinoleic and eleostearic acids, respectively (Kroon et al., 2006; Shockey et al.,

2006). These plant DGAT2s were functionally characterized in yeast cells and found to have a preference for substrates containing these uncommon FAs.

In addition to DGAT1 and DGAT2, several other enzymes were reported to catalyze the synthesis of TAG. A bifunctional WS/DGAT, possessing both enzyme activities, was identified in the bacterium *Acinetobacter* (Kalscheuer and Steinbuchel, 2003; Stoveken et al., 2005). The WS/DGAT was unrelated to known acyltransferases and DGATs from eukaryotes. Later, another WS/DGAT was identified in *A. thaliana*, which predominantly catalyzes the biosynthesis of wax esters (Li et al., 2008). Moreover, a cytosolic DGAT, which is non-homologous to either DGAT1 or DGAT2, was identified in developing peanut (*Arachis hypogaea*) cotyledons and referred to as AhDGAT or DGAT3 (Saha et al., 2006). Another soluble DGAT was also characterized in *A. thaliana*, which shares sequence homology with other plant DGAT1s and is involved in catalyzing the assembly of hydroxy FA-containing TAG which may serve as a precursor for cutin synthesis (Rani et al., 2010). Previously, a cytosolic 10S multienzyme complex in oleaginous yeast was found to possess DGAT activity (Gangar et al., 2001). Taken together, these studies provide direct evidence for the existence of TAG biosynthesis in cytosol. Soluble forms of DGAT, however, are in need to be characterized more thoroughly in a range of species.

Currently, a relatively large collection of DGAT genes is available, facilitating more detailed studies on enzyme structure and function. Many of these genes have been functionally characterized in recombinant systems as summarized in Table 2.1.

Table 2.1. Acyl-CoA:diacylglycerol acyltransferases (DGATs) functionally characterized in recombinant organisms (based on Siloto et al., 2009a). The two-character code preceding each *DGAT* gene indicates the organisms of origin are as follows: At, *Arabidoposis thaliana*; Ah, *Arachis hypogaea*; Bn, *Brassica napus*; Ce, *Caenorhabditis elegans*; Ea, *Euonymus alatus*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Nt, *Nicotiana tabacum*; Rc, *Ricinus communis*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Tg, *Toxoplasma gondii*; Tm, *Tropaeolum majus*; Ur, *Umbelopsis ramanniana*; Vf, *Vernicia fordii*; Vg, *Vernonia galamensis*; Zm, *Zea mays*. Reprinted with permission of John Wiley & Sons, Ltd.

Gene	cDNA	Host used for expression and relevant genetic markers	Reference
	AtDGAT1	Sf21 insect cells (*Spodoptera frugiperda*)	Hobbs et al., 1999
		S. cerevisiae YMN5 (*Slc1-Δ*)	Zou et al., 1999
		S. cerevisiae SCY 062	Bouvier-Nave et al., 2000a
		S. cerevisiae SCY059 (*are1-Δ, are2-Δ*), *N. tabacum*	Bouvier-Nave et al., 2000b
		A. thaliana AS11	Jako et al., 2001
		S. cerevisiae H1266 (*are2-Δ, dga1-Δ, lro1-Δ*)	Milcamps et al., 2005
		B. napus	Weselake et al., 2008
	BnDGAT1	*P. pastoris*	Nykiforuk et al., 2002
		B. napus	Weselake et al., 2008
	TmDGAT1	*S. cerevisiae* H1246 (*are1-Δ, are2-Δ, dga1-Δ, lro1-Δ*), *A. thaliana* and *B. napus*	Xu et al., 2008
DGAT1	*ZmDGAT1*	*Z. mays* and *S. cerevisiae* (*dga1-Δ, lro1-Δ*)	Zheng et al., 2008
	VgDGAT1	*S. cerevisiae*	Yu et al., 2008
	VfDGAT1	*S. cerevisiae* SCY1998 (*dga1-Δ, lro1-Δ*)	Shockey et al., 2006
	RcDGAT1	*S. cerevisiae*	He et al., 2004
	TgDGAT1	*S. cerevisiae* SCY910 (*are1-Δ, are2-Δ*)	Quittnat et al., 2004
	NtDGAT1	*S. cerevisiae* SCY 062	Bouvier-Nave et al., 2000a
	HsDGAT1	McA-RH7777 cells (*Rattus norvegicus*)	Liang et al., 2004
		S. cerevisiae 12501 (*dga1-Δ*)	Inokoshi et al., 2009
	MmDGAT1	Sf9 insect cells (*S. frugiperda*)	Cases et al., 2001
		COS7-cells (*Chlorocebus sabaeus*)	Yen et al., 2005
		H5 insect cells (*Trichoplusia ni*)	Cases et al., 1998
	EaDGAT1	*S. cerevisiae* H1266 (*are2-Δ, dga1-Δ, lro1-Δ*)	Milcamps et al., 2005
	CeDGAT1	*S. cerevisiae* SCY 062	Bouvier-Nave et al., 2000a
	HsDGAT2	*Sf9 insect cells (S. frugiperda)*	Cases et al., 2001
		S. cerevisiae ScY2051 (are2-Δ, dga1-Δ, lro1-Δ)	Turkish et al., 2005
		S. cerevisiae 12501 (dga1-Δ)	Inokoshi et al., 2009
	MmDGAT2	*Sf9 insect cells (S. frugiperda)*	Cases et al., 2001
		COS7-cells (C. sabaeus)	Stone et al., 2006
DGAT2	*RcDGAT2*	*S. cerevisiae*	Kroon et al., 2006
		A. thaliana and S. cerevisiae (dga1-Δ)	Burgal et al., 2008
	VfDGAT2	*S. cerevisiae SCY1998 (dga1-Δ,lro1-Δ)*	Shockey et al., 2006
	CeDGAT2	*Sf9 insect cells (S. frugiperda)*	Lardizabal et al., 2001

Table 2.1 Continued

	ScDGAT2	*Sf9 insect cells (S. frugiperda)*	Lardizabal et al., 2001
	SpDGAT2	*S. pombe DKO (dga1-Δ,lro1-Δ)*	Zhang et al., 2003
DGAT2	*UrDGAT2*	*Sf9 insect cells (S. frugiperda)*	Lardizabal et al., 2001
		G. max	Lardizabal et al., 2008
Soluble	*AhDGAT*	*E.coli BL21 (DE3)*	Saha et al., 2006
DGAT	*AtDGAT*	*E.coli BL21 (DE3)*	Rani et al., 2010

2.3 Functional and structural features of DGAT1 and DGAT2

Both DGAT1 and DGAT2 function as DGAT enzymes as demonstrated by forward and reverse genetics studies, but they are not similar at the level of DNA or amino acid sequences (Yen et al., 2008). Here, the structural and functional aspects of these two unrelated proteins are summarized with particular attention to gene structure and expression pattern, protein primary structure, biochemical properties, membrane topology and subcellular localization.

2.3.1 Gene organization and expression pattern

The structure of genes encoding DGAT1 and DGAT2 has been analyzed by Siloto et al. (2009a) based on genome database searching. Generally, *DGAT1* comprises 16-17 exons whilst *DGAT2* contains only 8-10 exons. Different gene orthologues within each *DGAT* family share similar architectures in the same kingdom of organisms except for the invertebrate *Caenorhabditis elegans*. In this organism, *DGAT1* and *DGAT2* show an unrelated distribution to other animals and contain fewer exons compared to other eukaryotes (Figure 2.2). It is noted that in animals, the exons are grouped in the 3' end while in plants they are distributed throughout the whole sequence.

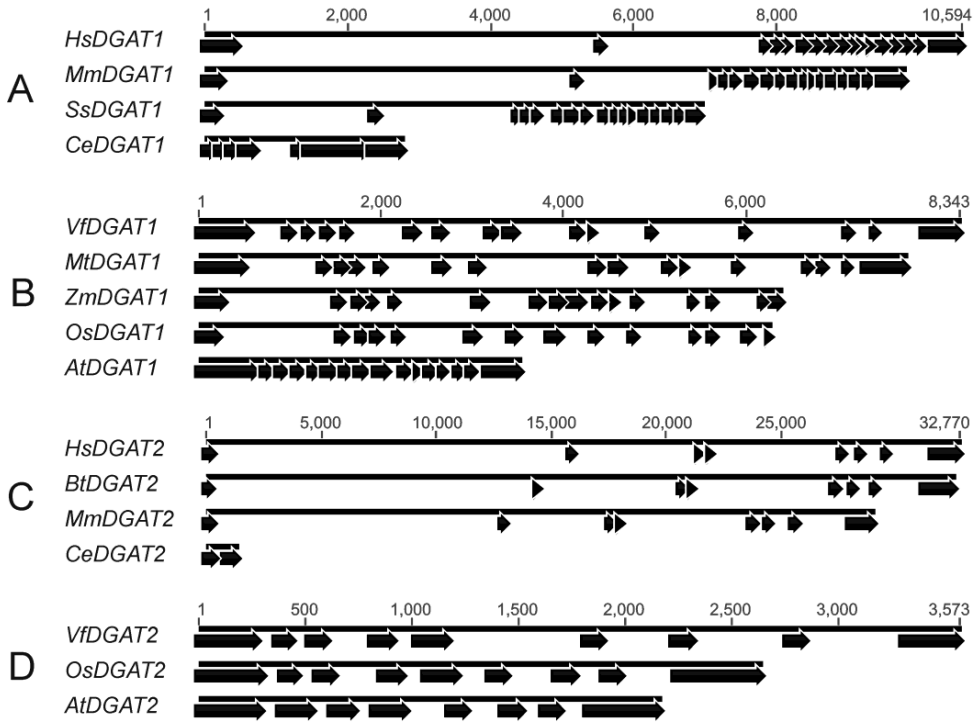


Figure 2.2. Architecture of representative acyl-CoA:diacylglycerol acyltransferase (*DGAT*) genes (Siloto et al., 2009a). (A) *DGAT1* from animals, (B) *DGAT1* from plants, (C) *DGAT2* from animals and (D) *DGAT2* from plants. The genomic sequences of each DGAT are represented by black bars. The arrows correspond to the regions comprising the coding region of the mRNA. The numbers correspond to the nucleotide positions. The two-character codes preceding each *DGAT* gene indicates the organism of origin are as follows: At, *Arabidoposis thaliana*; Bt, *Bos Taurus*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Mt, *Medicago truncatula*; Os, *Oryza sativa*; Sc, *Saccharomyces cerevisiae*; Ss, *Sus scrofa*; Vf, *Vernicia fordii*; Zm, *Zea mays*. Reprinted with permission of John Wiley & Sons, Ltd.

Besides the differences in gene organization, *DGAT1* and *DGAT2* have significant dissimilarities in expression pattern. In mammals, such as humans and mouse, *DGAT1* is ubiquitously expressed with the highest level in small intestine, while *DGAT2* expression is most abundant in liver and adipose tissue (Cases et al., 1998; 2001). *DGAT* expression has been more extensively studied in developing seeds of plants. Generally, the expression level of *DGAT1* or *DGAT2* is closely correlated with oil deposition in developing seeds of oil-bearing plants; however, the relative expression of *DGAT1* versus *DGAT2* varies depending on the species. In common oilseed crops, *DGAT1* has higher expression levels than *DGAT2*, indicating it plays a major role in seed oil production in these species (Li et al., 2010). But in plants accumulating unusual FAs such as tung tree and castor bean, *DGAT2* was expressed at higher levels than *DGAT1* in seeds (Kroon et al., 2006; Shockey et al., 2006). During seed development in these plants, *DGAT1* expression was slightly elevated whereas *DGAT2* expression was strongly boosted before the onset of oil biosynthesis. This is in line with biochemical work showing that DGAT2 from these species had a preference for unusual substrates (Kroon et al., 2006; Shockey et al., 2006). DGAT2 appears to be essential for the selective accumulation of these specific FAs in TAG. Expression of *DGAT*s was also detected in many tissues other than seeds (Hobbs et al., 1999; Xu et al., 2008) suggesting that these enzymes may be related to other physiological processes in

addition to seed oil synthesis. For example, *DGAT1* has been shown to be highly expressed during pollen development, presumably accounting for TAG accumulation in the pollen grain (Lu et al., 2003; Zhang et al., 2009).

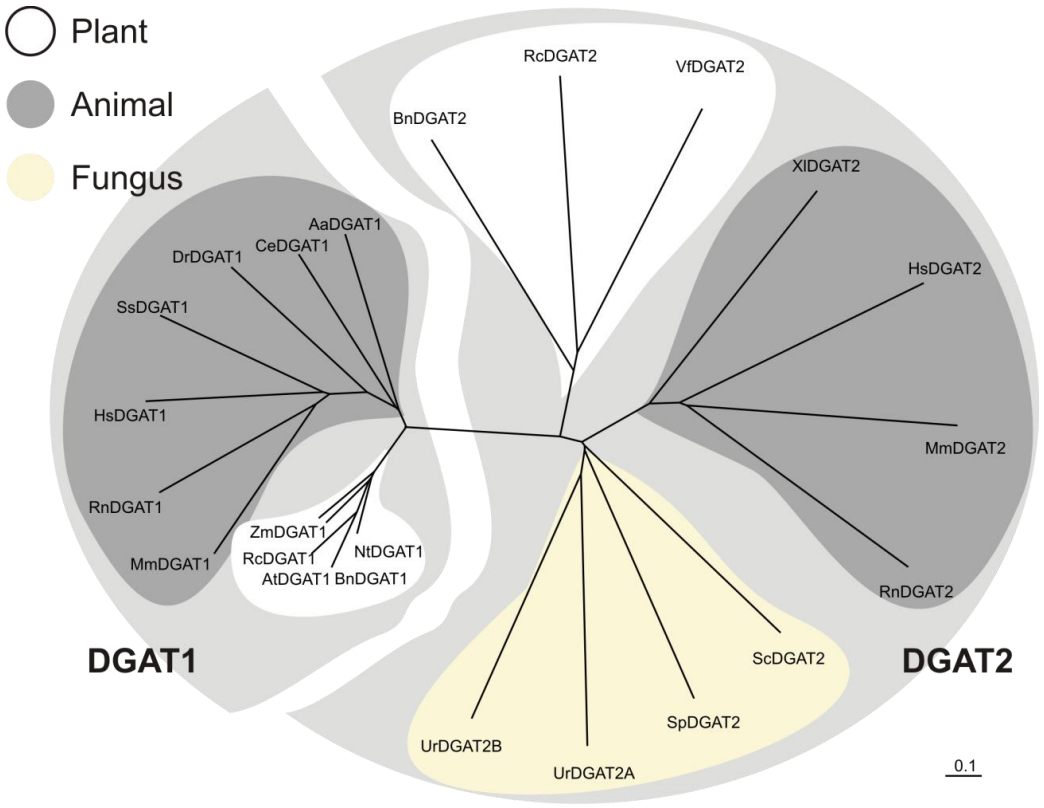
2.3.2 Protein primary structures

In most organisms, *DGAT1* encodes a protein of around 500 amino acid residues with a molecular mass of about 55kDa. Hydropathy plot analysis indicates that DGAT1 is typically characterized by a hydrophilic N-terminus followed by a number of hydrophobic stretches (Shockey et al., 2006; Siloto et al., 2009a). In contrast, DGAT2 is generally less than 400 amino acid residues long, less hydrophobic and has a molecular mass of about 40kDa (Shockey et al., 2006; Siloto et al., 2009a). A phylogenetic analysis of representative sequences of DGAT1 and DGAT2 clearly demonstrated a complete lack of similarity between these two classes of proteins (Figure 2.3). In each kingdom, the resulting cladogram contains two main divisions, made up entirely of either DGAT1 or DGAT2 sequences, except in most of fungi which contains only DGAT2.

The DGAT1 amino acid sequence is 15-25% identical to ACATs with the most conserved region in the C termini. DGAT1 and ACAT belong to the superfamily of membrane-bound O-acyltransferase (MBOAT, National Center for Biological Information (NCBI) conserved Domains Database accession number: pfam03062), which catalyzes the transfer of acyl chains onto hydroxyl or thiol

groups of lipids and proteins (Hofmann, 2000). The MBOAT family is characterized by a highly conserved histidine (H) within a long hydrophobic region. This residue was proposed to be involved in the active site. This hypothesis is supported by experimental evidence showing that this conserved histidine residue is essential for activities of both human ACAT1 and murine DGAT1 (Guo et al., 2005b; McFie et al., 2010). Previously, a conserved motif containing a histidine residue was also found in two other acyltransferases, GPAT and LPAAT, which are also involved in TAG bioassembly (section 2.1). Site-directed mutagenesis studies suggested that this motif is an important part of the active site (Frentzen, 1998). The histidine residue was hypothesized to abstract a proton from the hydroxy group of the acyl acceptor (G3P or LPA), facilitating nucleophilic attack on the thioester bond of the acyl donor (acyl-CoA). It was anticipated that the histidine residue in the MBOAT motif could function by a similar mechanism. A conserved histidine residue in DGAT2 was also demonstrated to be essential for enzyme catalysis (Stone et al., 2006). Nevertheless, phylogenetic analysis has shown a complete lack of similarity between DGAT2 and MBOAT members (Hishikawa et al., 2008) suggesting that DGAT2 may not have a similar catalytic mechanism to MBOAT enzymes.

Figure 2.3. Phylogenetic analysis of representative DGAT1 and DGAT2 sequences from plants, humans, and fungi. The branch lengths of the tree are proportional to divergence. The 0.1 scale represents 10% change. The two-character code preceding each DGAT indicates the organism of origin as described as follows: Aa, *Aedes aegypti*; Bn, *Brassica napus*; Ce, *Caenorhabditis elegans*; Dr, *Danio rerio*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Nt, *Nicotiana tabacum*; Rc, *Ricinus communis*; Rn, *Rattus norvegicus*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ss, *Sus scrofa*; Ur, *Umbelopsis ramanniana*, Vf, *Vernicia fordii*; XL, *Xenopus laevis*; Zm, *Zea mays*.


Plant
Animal
Fungus
RcDGAT2
VfDGAT2
BnDGAT2
XlDGAT2
HsDGAT2
AaDGAT1
CeDGAT1
DrDGAT1
SsDGAT1
HsDGAT1
MmDGAT2
RnDGAT1
ZmDGAT1
NtDGAT1
RcDGAT1
MmDGAT1
AtDGAT1
BnDGAT1
RnDGAT2
DGAT1
DGAT2
ScDGAT2
SpDGAT2
UrDGAT2B
UrDGAT2A
0.1

DGAT2 has recently been proposed to belong to another family (DAGAT, NCBI Conserved Domain Database accession number: pfam03982), which contains acyl-CoA:monoacylglycerol acyltransferase (MGAT) and acyl-CoA:wax alcohol acyltransferase (AWAT) (Yen et al., 2008). Three MGAT isoforms including MGAT1, MGAT2 and MGAT3 were identified, which catalyze DAG formation in insect and/or mammalian cells (Cheng et al., 2003; Yen and Farese, 2003; Yen et al., 2002). Two additional members, AWAT1 and AWAT2, mediating synthesis of wax esters, were also found to be capable of esterifying DAG in yeast cells (Turkish et al., 2005). The amino acid sequences of MGATs and AWATs shared significant homology with yeast and human DGAT2 proteins, with marked conservation of sequence in the C terminal regions. To date, MGAT and AWAT orthologues have only been identified in humans and mice. The functional significance of these enzymes remains to be determined.

2.3.3 Biochemical properties

The biochemical properties of DGAT have been explored largely through *in vitro* enzyme assays involving production of radiolabeled TAG from radiolabeled acyl-CoA and unlabeled *sn*-1,2-diacylglycerol (DAG) catalyzed by microsomal fractions (Lung and Weselake, 2006). Because the hydrophobic nature of DGAT made the purification difficult, microsomal fractions are typically used to quantify DGAT activity. Microsomal proteins from higher

eukaryotes, however, are likely to contain both DGAT1 and DGAT2 polypeptides. Therefore, it was impossible to ascribe substrate specificity to one specific DGAT isoform in the earlier assays with microsomal fractions (Lung and Weselake, 2006). DGAT1 and DGAT2 have been functionally characterized at the molecular level leading to new molecular tools for studying TAG biosynthesis. It is now feasible, for example, to express a specific *DGAT* cDNA in a cell devoid of endogenous DGAT activity, thus allowing for a more effective characterization of specific biochemical properties. As a result, dissimilarity in biochemical properties between DGAT1 and DGAT2 has been revealed in different organisms.

Different biochemical properties of two DGATs were first demonstrated in a study with mammalian orthologues. Cases et al. (2001) reported that recombinant DGAT1 or DGAT2 from human or mouse have broad fatty acyl-CoA substrate specificity. In addition, DGAT1 preferred a monounsaturated substrate, oleoyl (18:1 *cis*Δ9)-CoA as compared with saturated palmitoyl (16:0)-CoA. Murine DGAT1 also possessed additional enzyme activities such as MGAT and acyl-CoA:retinol acyltransferase (ARAT) activities (Yen et al., 2005). Since the DGAT2 family contains MGATs, it is attempting to speculate that DGAT1 and DGAT2 may have some overlapping functions, at least in mammals. Another possibility is that murine DGAT1 has a less discriminating substrate binding site

than DGAT2 (Yen et al., 2008). These hypotheses still need to be validated through further experimentation. In addition, human and mouse DGAT2s were proven to be highly active DGATs with a higher affinity for their substrate than DGAT1 (Cases et al., 2001). In rat hepatoma cells, *DGAT2* overexpression led to considerably more TAG accumulation than *DGAT1* overexpression (Stone et al., 2004), although DGAT2 activity was relatively lower *in vitro*. This might suggest that the assay conditions for DGAT2 are different from DGAT1. It could also indicate that effective DGAT catalysis, particularly for DGAT2, may involve unknown cofactors which are not present in the *in vitro* assays because they were lost during the preparation of microsomes.

The substrate preference of DGAT2 orthologs display differences across different species of fungi. DGAT2 from *U. ramanniana* showed enhanced activity towards medium-chain fatty acyl-CoAs (e.g, 12:0-CoA) and had enhanced preference for short- and medium-chain DAG substrates (6:0, 8:0 and 10:0) (Lardizabal et al., 2001). Yeast *S. cerevisiae* DGAT2 (Dga1p) used 18:1-CoA (18:1) and 16:0-CoA more efficiently than 14:0-CoA , 18:0-CoA (18:0), arachidonyl (20:4 *cis*Δ5,8,11,14)-CoA, and linoleoyl (18:2 *cis*Δ9,12)-CoA (Oelkers et al., 2002). In *S. pombe*, DGAT2 prefers palmitoyl over oleoyl moieties (Zhang et al., 2003). Recently, in *Claviceps purpurea*, which produces TAGs rich in ricinoleic acid, DGAT2 was demonstrated to prefer hydroxyl FA

ricinoleic acid (12-hydroxy 18:1, *cis*Δ 9) as an acyl donor over linoleic acid (18:2), oleic acid (18:1) or α-linolenic acid (18:3 *cis* Δ9,12,15) (Mavraganis et al., 2010). Thus, *C. purpurea* DGAT2 could be an attractive candidate gene for engineering of oilseed crops to produce high levels of hydroxy FAs.

A comparative analysis of plant DGAT1 and DGAT2 has been performed in tung tree (*V. fordii*). Tung DGAT2, but not DGAT1, was suggested to have a propensity for α-eleostearic acids (18:3, *cis*Δ9, *trans*Δ11, *trans*Δ13) rich in tung oil by expression pattern analysis in section 2.3.1, but the results from the *in vitro* enzyme assay did not show clear differences between DGAT1 and DGAT2 regarding the substrate preference (Shockey et al., 2006). A novel functional complementation assay to investigate the enzyme preference of tung DGATs *in vivo* was also applied in this study. In this assay, *wild type* or mutant yeast cells were cultivated in the presence of exogenous tung oil and a nonspecific lipase from *Candida rugosa*. Interestingly, although tung DGAT2 had a lower activity than DGAT1 in *in vitro* assay, it possessed approximately five-fold preference for incorporation of α-eleostearic acid into trieleostearin than DGAT1 *in vivo*. As for the DGAT2 preference for specific FAs, castor bean DGAT2 was also found to prefer hydroxy FA substrates such as ricinoleic acid, the major constituent of castor bean oil (Burgal et al., 2008; Kroon et al., 2006). Taken together with the results from the expression pattern study (Li et al., 2010), these data provide

additional evidence which supports the hypothesis that DGAT2 may play a role in the selective accumulation of unusual FAs in the seed oil of some plants (Li et al., 2010). As noted above, the fungal DGAT2 from *C. purpurea* also prefers ricinoleic acid, indicating that in certain fungal species, DGAT2 also has a preference to incorporate unusual FAs into TAG.

In addition to the substrate preference of DGATs, DGAT modulators including activators and inhibitors have also been investigated. Certain proteins and electrolyte have been shown to modulate microsomal DGAT activity (Lung and Weselake, 2006). Nonetheless, microsomes utilized in these studies were isolated from sources which possibly had more than one isoform of DGAT (Lung and Weselake, 2006). Thus, our understanding of the molecular mechanism by which activators or inhibitors exert their effects on specific DGAT isoforms is limited. This may be overcome by studying individual DGAT isoforms in yeast devoid of endogenous TAG synthesis (Sandager et al., 2002). For example, acyl-CoA binding protein (ACBP) has been shown to modulate DGAT utilization of acyl-CoAs depending on the ratio of ACBP to acyl-CoA in the reaction mixture (Yurchenko, 2009). Also, bovine serum albumin (BSA), had a similarly limited stimulatory effect on *B. napus DGAT1* (Yurchenko, 2009). This data on the effect of BSA differed from earlier reports in *A. thaliana* and *B. napus* in which DGAT activities were substantially increased with the presence of BSA (Hobbs and Hills,

2000; Little et al., 1994). In addition, it has also been shown that microsomal DGAT activity from muscle tissue of *Bos taurus* is inhibited by the thiol-modifying reagent *N*-ethylmaleimide (NEM) (Lozeman et al., 2001). This finding is in agreement with previous reports in rat myotube tissue and the fungus, *U. ramanniana* (Kamisaka et al., 1993; Sauro and Strickland, 1990). To further validate and explore the role of cysteine residues in the thiol-specific inhibition of yeast DGAT2, detailed function and structure studies have been performed which will be discussed in chapter 4.

2.3.4 Membrane topology

2.3.4.1 Topological characterization of DGATs

Understanding the topological pattern in which a protein transverses the membrane bilayer is essential for elucidating the catalytic mechanism of a membrane-bound protein. Both DGAT1 and DGAT2 contain hydrophobic segments which are generally believed to constitute transmembrane domains (TMDs) (Figure 2.4). DGAT1 is more hydrophobic than DGAT2 and is predicted to contain multiple TMDs. Initial work from tung tree DGATs has demonstrated that both termini of DGAT1 face the cytosol, supporting a proposed ten-TMD model based on *in silico* analysis (Shockey et al., 2006). Recently, experimental data from McFie et al. (2010) revealed a topological model of murine DGAT1 which contained only three TMDs with the N-terminus oriented towards the

cytosol. Moreover, the C-terminal region resides on the luminal side of the ER and probably contains the active site with the highly conserved histidine residue residing in the MBOAT motif as discussed in section 2.3.2 (His^{426} in murine DGAT1). It has been speculated that the different topological organization of mammalian and plant DGAT1s may due to the low level of sequence homology in certain regions (McFie et al., 2010). This type of discrepancy in structural features between protein orthologues from different categories of organisms was also observed in our DGAT2 studies discussed in Chapter 5.

Although DGAT2 is less hydrophobic compared to DGAT1 (Figure 2.4) with a less intricate topology, murine DGAT2 is the only DGAT2 member for which a topology model has been experimentally determined (Stone et al., 2006). In this model, one or two adjacent TMDs reside at the N-terminus of the protein with both termini residing in the cytosol. The bulk of the C-terminal region is exposed to the cytosol and contains the possible active site, a highly conserved HPHG motif (residues 161-164 of murine DGAT2). This model is consistent with earlier work in determining the N/C termini sidedness of tung tree DGAT2 (Shockey et al., 2006).

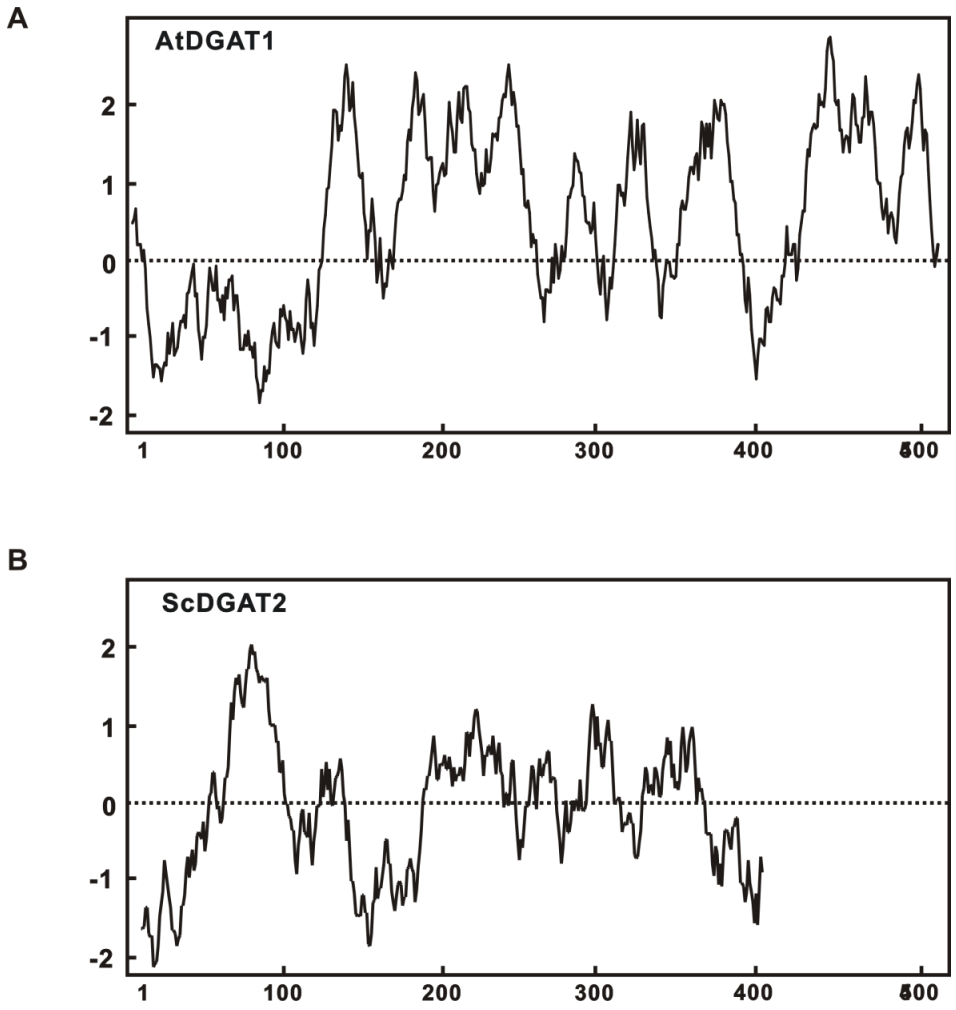


Figure 2.4. Kyte-Doolittle hydropathy plots of DGATs. Representative DGAT sequences including AtDGAT1 (A) and ScDGAT2 (B) were used. Plots were generated by the method of Kyte and Doolittle (1982) using a window size of 19. The two-character code preceding each DGAT polypeptide indicates the organism of origin as follows: At, *Arabidoposis thaliana*; Sc, *Saccharomyces cerevisiae*.

Nonetheless, topological studies of DGAT1s have shown that the N/C terminus orientation, with respect to cytosol/lumen, is different in tung tree versus murine DGAT1. Therefore, the topological orientation of DGAT1 orhtologues is not universally conserved across different species. It is also possible that a similar situation holds for DGAT2 from various species. Since murine DGAT2 is the only published topology for a DGAT2 orthologue, it would appear prudent to investigate the topological organization of other DGAT2.

2.3.4.2 Experimental approaches for determing membrane topology

Generally, to determine the topological organization of a membrane protein, a model will be first predicted using a variety of prediction programs based on the hydropathy plot method initiated by Kyte and Doolittle (1982). Most of the prediction algorithms are developed for α-helical membrane proteins since only a few membrane proteins from bacteria are so far known to possess β-barrel secondary structure. The techniques used by these programs can be summarized in six types: hydrophobicity analysis (DAS), positive inside rule (TopPred), multiple sequence alignment (TMAP), model-recognition approach (MEMSTAT3), support vector machine technique (SVMtm) and consensus techniques (Conpred II) (Persson, 2006). Topology models generated by different prediction methods, however, can be different. Thus, multiple algorithms are routinely used to survey possible topological features as prelude to experimental

verification.

Experimental strategies for topology assessment can be divided into two approaches (van Geest and Lolkema, 2000). One earlier and still used approach is to make a gene fusion in which a reporter enzyme is attached to different hydrophilic portions of a protein (Boyd, 1996). Because the properties of the enzyme (e.g. enzymatic activity) depend on their subcellular location, one can determine at which side of the membrane the fusion site resides. Commonly used reporter enzymes include alkaline phosphatase and β-galactosidase (Bogdanov et al., 2005). Typically, the large reporter enzyme is fused to the truncated C-terminus of the membrane protein or inserted into the membrane protein, which could possibly disrupt the native topology of the membrane protein. Thus, this method has only be used with success for certain prokaryote membrane proteins (van Geest and Lolkema, 2000).

Another more generally used strategy is to analyze the accessibility of certain introduced or endogenous tags to specific reagents with different membrane permeabilities. The introduced/endogenous tags commonly used contain N-glycosylation sites (Gilstring and Ljungdahl, 2000), antibody epitopes (Stone et al., 2006), proteolytic sites (Gilstring and Ljungdahl, 2000) or cysteine residues (Guo et al., 2005b). By determining the accessibility of the tag at different positions in the protein, the complete topology can be characterized. In

this study, cysteine residue was chosen as the preferred tag because of the following strengths (Bogdanov et al., 2005; van Geest and Lolkema, 2000). Substituon of target residue with cysteine residue are generally well tolerated with respect to maintain the native topology and protein function. Modification by chemical reagents is restricted to single residue in a complete and active protein. Moreover, with the availability of different commercial cysteine-specific reagents, modification with denatured protein could generate more structural information than topological orientation alone, for example, mapping of disulfide linkageis as demonstrated in Chapter 4 and in previous work on ACAT (Guo et al., 2005a). Also, in the case of ScDGAT2, an active mutant protein devoid of endogenous cysteine was prepared for generating single cysteine-containing mutants (detailed in Chapter 4). An intrinsic limitation of this approach, however, is that introduced cysteine residues at certain positions close to TMD, or buried in highly structured regions, could have limited accessibility to chemical modification, which could affect correct assignment of location to these positions. In our topology work described in Chapter 5, we have encountered this situation and the related issue has been detailed and addressed in Chapter 6.

2.3.5 Subcellular localization

As described previously, the ER is regarded to be the main site for TAG synthesis (Lung and Weselake, 2006). To better elucidate the role of DGATs in

this cellular process, the subcellular location of DGAT1 and DGAT2 has been studied in various organisms including mammals, plants and fungi.

In tung tree, DGAT1 and DGAT2 have been demonstrated to localize in the ER due to a C-terminal ER retrieval motif (Shockey et al., 2006). The ER retrieval sequence in tung DGAT1 is YYHDL, part of the motif LLYYHDXMN, conserved in DGAT1 from plants. In tung DGAT2, the ER retrieval motif comprises the sequence LKLEI, in which the two leucines are conserved in other plant DGAT2 sequences. Although both DGATs were shown to be targeted to the ER, DGAT1 and DGAT2 were shown to be located in separate regions of this membrane network. Taken together with the mounting evidence from the genetic and biochemical studies of tung DGATs discussed previously, it was speculated that these two DGATs may have distinct interaction with other proteins in the ER as discrete multiprotein complexes (Yen et al., 2008). DGAT1 activity was also identified in the chloroplasts of senescing Arabidopsis leaves (Kaup et al., 2002). The mechanism by which DGAT1 transported to chloroplast is yet unclear.

In fungi and animals, in addition to the ER, DGAT activity has also been reported to be present in cystosolic lipid droplets, the major site for storing intracellular TAG. For example, in yeast (*S. cerevisiae*), DGAT activity (DGAT2) has been reported in lipid droplets (Sorger and Daum, 2002). This was supported by two subcellular localization datasets generated by proteomic analysis of *S.*

cerevisiae (Huh et al., 2003; Natter et al., 2005). In addition, fluorescence imaging of TAG biosynthesis in mammalian cells showed human DGAT2 was also associated with lipid droplets (Kuerschner et al., 2008). This ER/lipid droplet dual localization was further investigated by detailed localization studies on murine DGAT1 and DGAT2 expressed in cultured mammalian cells (Stone et al., 2009). Murine DGAT2 was found to localize to ER in cells without treatment with exogenous FAs. In contrast, when exogenous FAs were provided, DGAT2 also co-localized to mitochondrial-associated membranes (MAM). MAM is a membrane bridge between the ER and mitochondria, which is near the surface of lipid droplets. Although an N-terminal mitochondrial targeting signal sequence was identified, it is not clear why DGAT2 is also associated with lipid droplets. There were two proposed possibilities for the ways by which the murine DGAT2 is affiliated with lipid droplets (Stone et al., 2009; Yen et al., 2008). One was that the N-terminal TMDs in murine DGAT2 were directly embedded into the surface of the lipid droplet similar to the way of oleosin binds to the surface of oil bodies in plant cells (Siloto et al., 2009a). Oleosin has a long hydrophobic domain composed of two chains rooted into the lipid bodies and two cystosolic termini (Abell et al., 2004; Tzen et al., 1992). Similarly, murine DGAT2 contains two putative adjacent TMDs with both termini facing the cytosolic side. A second possibility is that the TMDs of murine DGAT2 in the ER membrane bilayer are in

close proximity to lipid droplets (Stone et al., 2009; Yen et al., 2008). The latter hypothesis was favored because proteomics studies of isolated lipid droplets have not revealed the presence of DGAT1 or DGAT2 polypeptides (Stone et al., 2009; Yen et al., 2008). In addition, the length of the hydrophobic domain of murine DGAT2 may not be long enough to allow the polypeptide to penetrate into the oil body membrane (Herman, 2008). It is noteworthy that murine DGAT1 did not localize to other sites other than ER in the cells with or without FA treatment (Cases et al., 2001). Collectively, these studies suggest that murine DGAT1 and DGAT2 have distinct localization patterns and interaction with other proteins, similar to observations in tung tree. Supporting this, murine DGAT2 expressed in a human cell line was demonstrated to colocalize and physically interact with stearoyl-CoA desaturase 1 (SCD), an ER enzyme catalyzing the desaturation of acyl-CoAs to produce monounsaturated acyl moieties (Man et al., 2006). Therefore, in mammals, DGAT2 could be responsible for channeling endogenous monounsaturated FAs into TAG. It is not clear if DGAT1 is also involved since interaction of DGAT1 and SCD is yet to be investigated. A study to elucidate the role of DGAT2 in very low-density lipoprotein (VLDL) secretion provides additional evidence supporting the hypothesis in which DGAT1 and DGAT2 may localize in distinct regions of the ER. In this work, the presence or absence of

DGAT1 did not affect the process of DGAT2 being responsible for catalyzing the synthesis of TAG destined for secretion (Liu et al., 2008).

2.4 Biotechnological applications of DGAT1 and DGAT2

As discussed above, considerable advances have been occurred in the molecular and functional characterization of DGAT1 and DGAT2. The insights into the functional and structural aspects of DGATs help us to understand the functional significance of DGATs in TAG metabolism. In addition, insights into DGAT properties have prompted the development of DGAT applications in both therapeutics and oilseed engineering.

2.4.1 DGATs as therapeutic targets

During the process to discover different DGAT isoforms, gene knock-out studies in mouse models have laid a good foundation to examine the relationship between DGAT and diseases related to TAG metabolism. *DGAT1*-deficient mice (*Dgat1*$^{-/-}$), as described earlier, had reduced TAG levels in tissues but were healthy, fertile and were capable of TAG synthesis. Moreover, these lean mice were resistant to diet-induced obesity because of boosted energy disbursement and increased physical activity (Chen and Farese, 2000; Smith et al., 2000). Also, they had an enhanced sensitivity to both insulin and leptin (Chen et al., 2002a). These findings sparked considerable interest in exploring DGAT1 as a pivotal therapeutic target for treating obesity and diabetes in humans. Recently, a small

DGAT1 selective inhibitor was synthesized and reported to decrease weight and hepatic TAG level in an obesity-induced mouse model, which reproduced the phenotype of *DGAT1* knockout mice (*Dgat1*$^{-/-}$) (Zhao et al., 2008). King et al. (2009) further found that this potent and selective inhibitor could significantly reduce the serum TAG levels in both genetic and diet-induced mouse models of hypertriglyceridemia. Collectively, these compelling studies suggest that it is feasible to use DGAT1 inhibition as a novel therapeutic strategy to treat obesity, hepatic steatosis and hypertriglyceridemia. Different from *DGAT1*, *DGAT2* (*Dgat2*$^{-/-}$), knockout mice were lipopenic and had skin abnormalities, leading to an early death (Stone et al., 2004). Since *DGAT2* knockout is lethal for mice, from a safety perspective, inhibition of DGAT2 was initially less preferable as a potential therapy to treat TAG-related disorders (Yen et al., 2008). Recent studies in which antisense technology was used to disrupt *DGAT2* expression have proved this strategy is promising. Treatment with *DGAT2* gene-specific antisense oligonucleotide (ASO) to knock down *DGAT2* expression in mouse liver have resulted in a significant decrease of hepatic DGAT2 activity in the diet-induced mouse model of obesity (Liu et al., 2008). Moreover, TAG levels in both liver and plasma were substantially reduced without other side effects such as skin abnormalities (Yu et al., 2005). Also, this down-regulation of *DGAT* expression led to a remarkable decrease in body and adipose weight, hepatic TAG level and

insulin resistance induced by a high-fat diet (Choi et al., 2007; Liu et al., 2008). Interestingly, in the rat model of diet-induced NAFLD, suppression of *DGAT2*, but not *DGAT1*, with ASO treatment successfully improved hepatic steatosis and insulin sensitivity (Choi et al., 2007). Thus, DGAT2 inhibition also has a potential for treating hypertriglyceridemia, hepatic steatosis and obesity. In addition, *DGAT2* expression levels were significantly down-regulated in human psoriatic skin (Wakimoto et al., 2003). As described above, *DGAT2* knockout (*Dgat2*$^{-/-}$), in mice caused skin barrier abnormalities. Therefore, DGAT2 may also be a promising target for treating such diseases. Despite these findings, high-resolution structural and functional information of DGATs is still indispensible for understanding the mechanism by which the inhibitor or ASO limited the TAG overaccumulation in rodents. Eventually, this would aid in designing novel treatments beneficial for humans.

2.4.2 Use of DGATs in oilseed engineering

As crude oil resources decline, vegetable oils produced by oil-bearing crops are gaining growing interest as sustainable replacements for petroleum-derived chemical feedstocks for industrial applications. Thus, recent oilseed biotechnology has set out to expand the potential of increasing overall oil accumulation and producing desirable FA compositions in the seed oil. Both forward and reverse genetics have proven DGATs to be essential targets for such

applications as summarized in Table 2.2. For example, an EMS-induced mutation of *DGAT1* in *A. thaliana* (AS11) resulted in reduced seed oil content and delayed seed maturation (Zou et al., 1999). Similar findings were obtained from another *A. thaliana* mutant ABX45 caused by frame shift mutation of the gene (Routaboul et al., 1999). In addition, overexpression of *DGAT1* orthologues from plant sources including *A. thaliana* (Jako et al., 2001), *B.napus* (Weselake et al., 2008), and *Tropaeolum majus* (Xu et al., 2008) have been successfully used for boosting overall oil production in *A. thaliana* and *B. napus*. Also, Lardizabal et al. (2008) reported that overexpression of *DGAT2* from the fungus *U. ramanniana* also led to an increase of seed oil content in soybean (*Glycine max*), signifying that heterologous expression of fungal *DGAT* orthologues in plants could be also feasible for increasing overall oil production. Plant DGAT2 has been demonstrated to be selective for unusual FAs (Kroon et al., 2006; Shockey et al., 2006), providing important tools for producing industrially useful oils by coexpression of *DGAT2* and an enzyme responsible for the synthesis of an unusual FA from the same species in a transgenic host (Shockey et al., 2006). This hypothesis was subsequently tested in a study by Burgal et al. (2008) in which coexpression of FA hydroxylase (*FAH12*) and *DGAT2* from castor bean doubled the hydroxy FA content of the oil produced in transgenic *A. thaliana* compared to expression of *FAH12* alone. Recently, Mavraganis et al. (2010)

reported coexpression in yeast of *FAH* and ricinoleoyl-specific *DGAT2* from *C. purpurea*. As expected, their coexpression in yeast led to an elevated level of ricinoleic acid in the oil compared to expression of *FAH* alone. This finding suggests that fungal *DGAT2*s also represent attractive gene candidates for producing value-added novel oils in crops.

Table 2.2. Genetic engineering of oil-bearing crops through the use of DGAT (Based on Snyder et al., 2009). Reprinted with permission of Elsevier Science Ltd.

Host species	Source species	Up-regulated gene	Phenotype	Reference
Brassica napus	*B. napus*, *A. thaliana*	*DGAT1*	Increased seed oil content	Taylor et al., 2009; Weselake et al., 2008
A. thaliana, *N. tobaccum*	*A. thaliana*	*DGAT1*	Increased seed oil content, seed weight	Jako et al., 2001; Bouvier-Nave et al., 2000b
Glycine max	*U. ramanniana*	*DGAT2*	Increased seed oil content	Lardizabal et al., 2008
		FAH12		
A. thaliana	*R. communis*	*DGAT2*	Up to 30% hydroxyl fatty acids	Burgal et al., 2008
A. thaliana *Brassica napus*	*T. majus*	*DGAT1*	Increased seed oil content	Xu et al., 2008

It is noteworthy that DGAT1 could also play a role in channeling certain FAs into TAG. In *Arabidopsis*, knock out of *DGAT1* has also been shown to affect FA composition resulting a reduction in monounsaturated FAs (eg., 18:1 and 20:1) and enrichment in the polyunsaturated FA, α-18:3 (Katavic et al., 1995; Zou et al., 1999). This phenotype was reversed by restoring DGAT1 activity (Jako et al., 2001). Further understanding of the mechanism by which DGAT1 is involved in channeling these FAs into TAG fatty acids will likely provide additional insight into the malleability of TAG biosynthesis processes in oilseeds.

Overexpression of *DGAT* orthologues in transgenic plants has been successfully used to increase seed oil production with concomitant production of desirable FAs. In the longer term, approaches to enhance the overall catalytic efficiency and modulate substrate selectivity of DGATs are necessary to further raise the level of desired FA in the seed oil. Currently, molecular genetic approaches, such as SDM and directed evolution, are being explored as tools for these types of studies (Siloto et al., 2009b; Xu et al., 2008). Nevertheless, study on the relative contribution of DGAT1 and DGAT2 to seed oil accumulation in different plant species is still underway (discussed in Chapter 3). A detailed understanding of the catalytic mechanism and regulation of DGATs is important for further progress in metabolic engineering of TAG accumulation in oleaginous plants.

A version of part of this chapter has been published. Liu, Q., Siloto R. M. P. and Weselake R. J. 2010. Biochemistry. 49: 3237–3245.

3. Selection of a Acyl-CoA:Diacylglycerol Acyltransferase 2 for Structure/Function Studies

3.1 Introduction

Approximately 30% of genes in the sequenced genome encode for membrane proteins, which are an essential class of protein involved in many cellular functions and a wide variety of diseases (Gao and Cross, 2005). It is therefore not surprising that membrane proteins represent over 60% of the targets for existing drugs (Lundstrom, 2004). Despite their fundamental importance, the high-resolution structures of membrane proteins are determined at a low success rate. Currently, less than 2% of the structures deposited in the Protein Data Bank (http://www.wwpdb.org/) represent membrane proteins. The major bottleneck in structural determination of this class of protein is that it is often difficult to obtain quantitatively or qualitatively sufficient protein for structural analysis by such methods as crystallography and NMR spectroscopy (Little et al., 1994; McLuskey et al., 2010).

Acyl-CoA:diacylglycerol acyltransferase 2 (DGAT2), as an ER-associated protein, serves as a good example of this class of proteins due to its importance in TAG metabolism and as a target for treatment of diseases including

hypertriglyceridemia, hepatic steatosis and obesity (summarized in Chapter 2). So far, no three-dimensional structure has been reported for either DGAT1 or DGAT2. Although DGAT2 has been cloned for many years, only a truncated yeast DGAT2 has been purified to near homogeneity, with low quantities of protein obtained during the purification process (Kamisaka and Yasushi, 2010). With the development of recombinant DNA technology and the availability of numerous expression systems (Table 2.1, Chapter 2), molecular biology and biochemistry approaches have become powerful methods for exploring the structural and functional aspects of membrane proteins in recombinant form.

Before embarking on this type of study, it is necessary to select a recombinant DGAT2 that can be functionally expressed and which exhibits substantial biological activity in the expression system utilized. In this regard, the choice of expression system is also an important factor to consider. Attempts to produce eukaryotic membrane proteins in prokaryotic systems has often yielded poor results, with denatured protein accumulating in insoluble inclusion bodies (Reinhart and Krettler, 2006). *DGAT2* is a eukaryotic gene with orthologues functionally identified in various organisms including animals, fungi and plants (summarized in Chapter 2). Therefore, it is reasonable to choose a eukaryotic system *DGAT2* where potential post-translational modifications are more likely to occur for functional expression. As one of the predominant eukaryotic expression

systems, *Saccharomyces cerevisiae* (yeast) possesses many biotechnological advantages including ease of growth and handling, non-pathogenic characteristics and being an organism in which straightforward gene manipulation can be conducted (Romanos et al., 1992). Furthermore, *S. cerevisiae* has a well-characterized genetic system, and many different expression vectors and mutant strains are available for use in yeast. Thus, DGATs have often been characterized using the *S. cerevisiae* system combined with mutant strains (Siloto et al., 2009a).

The initial goal of this study was to identify a plant DGAT2 which could be functionally expressed in yeast, as a basis for continuing with structure/function studies on the plant enzyme. *DGAT2* cDNAs from two plant species and *S. cerevisiae* were examined for functional expression in an *S. cerevisiae* system which is deficient in endogenous DGAT activity and TAG synthesis due to the quadruple knockout of the four genes (*DGA1, LRO1, ARE1, and ARE2*) (Sandager et al., 2002). Of all the cDNAs tested, only *ScDGAT2* could be functionally expressed in the membrane fraction of yeast cells, presumably due to ineffective protein translation. The time course of ScDGAT2 production in yeast and the effect of epitope tag on the activity of the recombinant enzyme was also investigated, which provided a basis for the studies of structure/function which are the topics of Chapters 4 and 5.

3.2 Materials and methods

3.2.1 Cloning of *DGAT2* from *Arabidopsis thaliana*

A cDNA clone of *DGAT2 from A. thaliana (thale cress)* (pUNI 51-AtDGAT2) was obtained from *Arabidoposis* Biological Resource Center (Columbus, OH). The full-length DGAT2 open reading frame (ORF) was amplified by PCR employing the forward primer 5'-TATCGAATTCGCCATGGGTGGTTCTAGAGAGTTC-3' and the reverse primer 5'-GCGCCTCGAGAGAAGAATTTTCAGCTCAAG-3'. The PCR products were then purified and cloned into yeast expression vector pYES2.1-TOPO (Invitrogen) under the control of *GAL1* promoter and in frame with a V5 epitope tag, resulting in the AtDGAT2 polypeptide with an additional KGELRGHPFEGKPIPNPLLGLDSTRTGHHHHHH at the C-terminus. The constructs were verified by sequencing using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems)

3.2.2 Cloning of *DGAT2* from *Brassica napus*

A BLAST search of the protein sequence database (NCBI, NIH, Bethesda) using the AtDGAT2 (NP_566952) sequence revealed another two putative DGAT2 polypeptides from *Oryza sativa* (BAD33251, BAD07792). Several conserved regions of these three plant DGAT2 family members served as the starting points for isolation of *DGAT2* homologue from *B.napus (canola)*. Total

RNA was extracted from developing seeds of *B. napus* DH12075 line using the RNeasy Plant Mini Kit (QIAGEN, Ontario, Canada). The total RNA was then used to synthesize first strand cDNA using the SuperScript III First-Strand Synthesis System (Invitrogen). The synthesized cDNA was used as a template for PCR amplification with degenerate oligonucleotide primers designed according to the DGAT2 conserved motifs -(Q/H)IWTWLG- and -PMHVVG-, respectively.

The PCR products with expected size were purified and cloned into the pCR2.1-TOPO vector (Invitrogen) for sequencing. The sequence information obtained was used to design gene-specific primers for subsequent rapid amplification of cDNA ends (RACE) reactions. First-strand cDNA synthesized with adaptor-oligo dT (17) primer was used as the template for RACE reactions. 3' RACE reactions were completed using adapter primer 5'-GACTCGAGTCGACATCG-3' and gene-specific primer 5'-TCCAGTTTTCTCCTTCGGTCAGTCAGAT-3'. Initial attempts to isolate the remaining 5' cDNA ends by 5' RACE were unsuccessful. The terminal sequence of the 5' end was eventually identified by a BLAST search of the 3'RACE product sequence against an in-house *B.napus* expressed sequence tag (EST) database. The full ORF was then amplified using the forward primer 5'-TGATATGGGCAAAGTCAGAGAC-3' and the reverse primer 5'-AAGAATGTTCAACTGCAGATC-3' followed by cloning into pYES2.1-TOPO

(Invitrogen) for expression in *S. cerevisiae*. The cDNA sequences were deposited in the GenBank under the accession numbers FJ858270 and FJ858271, respectively.

3.2.3 Cloning of DGAT2 from *Saccharomyces cerevisiae*

Yeast genomic DNA was isolated from SCY62 (*MATα ADE2*) cells using the DNeasy Tissue Kit (QIAGEN). The ORF encoding DGAT2 from baker's yeast *S. cerevisiae* (*ScDGAT2*, YOR245C; designated DGA1) was PCR-amplified from yeast genomic DNA using the forward primer 5'-GCGATGTCAGGAACGTTCAATGATATAAG-3' and the reverse primer 5'-CCCAACTATCTTCAATTCTGCATCCGGTAC-3'. Following PCR amplification and gel purification, *ScDGAT2* was cloned into pYES2.1-TOPO (Invitrogen) as described in 3.2.1. An untagged DGAT2 was constructed as described above, except that the reverse primer introduced a stop codon (TAA) at the last codon.

3.2.4 Yeast strains and cell culturing

The *S. cerevisiae* stain with quadruple knockout of the *DGA1, LRO1, ARE1* and *ARE2* (H1246, *MATα are1*-Δ*::HIS3, are2*-Δ*::LEU2, dga1*-Δ*::KanMX4, lro1*-Δ*::TRP1 ADE2*) and the parental strain (SCY62, *MATα ADE2*) were supplied by Dr. S. Stymne and U. Ståhl (Swedish University of Agricultural Sciences) (Sandager et al., 2002). Wild type strain SCY62 was used for cloning of

ScDGAT2. H1246 was used as the host strain for protein expression experiments. *DGAT2* constructs were transformed into H1246 cells using lithium acetate/single-stranded carrier DNA/PEG method (Gietz and Schiestl, 2007). Yeast cells were also transformed with *pYESLacZ* or an empty plasmid as controls. The transformants were selected on plates lacking uracil and cultivated in minimal media containing 0.67% (w/v) yeast nitrogen base, 2% (w/v) dextrose, 20mg/L of adenine, arginine, tryptophan, methionine, histidine, and tyrosine, 30mg/L of lysine and 100mg/L of leucine. Cells were then harvested, washed with distilled water and inoculated in induction medium where dextrose was replaced by 2% (w/v) galactose and 1% (w/v) raffinose at OD_{600} of 0.4. The culture was incubated at 30^{o}C, 250 rpm to induce the expression of recombinant genes.

3.2.5 Gene expression analysis

Total RNA was extracted from yeast cells expressing *DGAT2* using RNA Mini Kit (QIAGEN). The first-strand cDNA was synthesized as described in 3.2.2 and used to amplify the target cDNA. The primers for specific amplification of *AtDGAT2* were 5'-TATCGAATTCGCCATGGGTGGTTCTAGAGAGTTC-3' (forward) and 5'-GCGCCTCGAGAGAAGAATTTTCAGCTCAAG-3' (reverse). The primers for *BnDGAT2s* amplification were 5'-TGATATGGGCAAAGTCAGAGAC-3' (forward) and 5'-

AAGAATGTTCAACTGCAGATC-3’ (reverse). The primers for amplification of *ScDGAT2* were 5’-GCGATGTCAGGAACGTTCAATGATATAAG-3’ (forward) and 5’-CCCAACTATCTTCAATTCTGCATCCGGTAC-3’ (reverse). PCR amplifications for *Actin* gene were performed as a control using primers 5’-ATGGATTCTGAGGTTGCTGCTTTGG-3’ (forward) and 5’-GAAACACTTGTGGTGAACGATAGATGG-3’ (reverse). PCR products were analyzed by agarose-gel electrophoresis.

3.2.6 Lipid extraction and analysis

The total lipid was extracted from yeast cells using a hexane-isopropanol method (Hara and Radin, 1978). An aliquot of the total lipid extract was spotted directly to silica gel thin layer chromatography (TLC) plate and separated using one ascension of n-hexane:diethyl ether:acetic acid (80:20:1, v/v). TAG spots were visualized by dipping the plates in a solution of 3% (w/v) cupric acetate in 8% (v/v) phosphoric acid and charring for 20mins.

3.2.7 DGAT assay

Induced yeast cells were harvested by centrifugation, washed with water and resuspended in 1mL cold extraction buffer (0.2M Hepes-NaOH, 0.5M sucrose, 1mM phenylmethylsulphonyl fluoride, pH 8.7). Yeast cells were lysed with 0.5mm glass beads in a bead beater cell homogenizer (Biospec, Bartlesville, USA). The lysate was centrifuged at 1500×*g* for 15min to remove the cell debris.

After a second centrifugation at 12,000×*g* for 10min, microsomes were recovered by ultracentrifugation at 100,000×*g* for 1h at 4ºC. The resulting pellets were resuspended in resuspension buffer 10mM Hepes-NaOH, pH 7.4. Equivalent amounts of protein, as determined by Bicinchoninic acid (BCA) assay (Pierce, Rockford, IL), were used for *in vitro* DGAT activity assays which were performed according to Weselake et al. (1991) with modifications. The reaction mixture contained 0.15M Hepes-NaOH (pH 7.0), 2.5mM $MgCl_2$, 0.1mg/mL bovine serum albumin (BSA, fatty acid free), 308μM *sn*-1,2-diolein, 15μM [1-^{14}C] oleoyl-CoA in a final volume of 55μL. The assay was initiated by 10μg of the protein in 10μL. After incubating for 10min at 30ºC, reactions were quenched by the addition of 10μL 10% SDS (w/v). Fifty microliters of the terminated reaction mixture were applied directly to silica gel TLC plates. Following extensive drying and application of 100μg of carrier triolein, TAG produced in the reaction mixture was resolved with n-hexane:diethyl ether:acetic acid (80:20:1, v/v). TAG spots were visualized with iodine vapor and scraped. Radioactivity was quantified by a Beckman-Coulter LS6500 liquid scintillation counter. Assays were performed in three biological replicates.

3.2.8 SDS-PAGE and immunoblotting

SDS-PAGE was conducted by adding loading buffer (125mM Tris, pH 6.8, 2% SDS, 25% glycerol, 5% 2-mercaptoethanol and 0.15% bromophenol blue) to

microsomal protein. Protein samples were separated in a 10% SDS-PAGE gel. Immunoblots were carried out by transferring proteins to a PVDF membrane. The recombinant protein was detected using anti-V5-HRP antibodies (Invitrogen) according to the protocol described by the manufacturer.

3.2.9 Phylogenetic analysis

Using AlignX (Invitrogen), amino acid sequences of representative *DGAT2* genes were aligned and the phylogenetic tree was constructed from the aligned sequences using the default algorithm. The GenBank accession numbers for genes used in the alignment are as follows: *A. thaliana*, At, NP_566952); *B. napus*, Bn, *BnDGAT2a*, FJ858270, *BnDGAT2b*, FJ858271; tung (*Vernicia fordii*), Vf, ABC94474; castor bean (*Ricinus communis*), Rc, AAY16324; human (*Homo sapiens*), Hs, AAK84176; Mouse (*Mus musculus*), Mm, AAK84175; *S. cerevisiae*, Sc, NP_014888; fission yeast (*Schizosaccharomyces pombe*), Sp, XP_001713160; *Umbelopsis ramanniana*, Ur, UrDGAT2a, AAK84179, UrDGAT2b, AAK84180.

3.3 Results

3.3.1 Identification of two cDNAs encoding DGAT2 in *B. napus*

To isolate the full-length coding sequence of *DGAT2* from *B. napus*, highly conserved sequences in other plant *DGAT2* genes were used to identify *DGAT2* cDNA in *B. napus* using standard homology-based cloning techniques

(detailed under “Materials and methods”). Two isoforms of full-length *BnDGAT2* cDNA were amplified and designated *BnDGAT2a* and *BnDGAT2b* (GenBank accession No. FJ858270 and FJ858271). The ORFs of both isoforms were 951 nucleotides in length with a high identity to each other (97.2%). Both isoforms encoded proteins of 317 amino acids in length with a predicted molecular mass of 35.6kDa. The deduced amino acid sequences of both isoforms share significant identity with all other representative and functional DGAT2s from major eukaryotic phyla (Table 3.1). BnDGAT2s are very similar to other plant DGAT2s with over 75% identity with *A. thaliana* DGAT2 (AtDGAT2) and over 50% identity to castor bean DGAT2 (*Ricinus communis*, RcDGAT2) or tung DGAT2 (*Vernicia fordii*, VfDGAT2). The BnDGAT2s also share around 20% identity with DGAT2s from animals and fungi. Preliminary phylogenetic analysis from the sequence alignment showed that BnDGAT2s are tightly clustered with other plant DGAT2s, which form a branch distant from two other groups of DGAT2s from animals and fungi (Figure 3.1).

Table 3.1. Amino acid identity between DGAT2 from *B. napus* and other functionally characterized DGAT2s. The two-character codes preceding each DGAT indicate the organisms of origin as follows: Hs, human (*Homo sapiens*); Mm, mice (*Mus musculus*); At (*A. thaliana*); Bn, *B. napus*; Vf, tung (*V. fordii*); Rc, castor bean (*R. communis*); Sc, budding yeast (*S. cerevisiae*); Sp, fission yeast (*Schizosaccharomyces pombe*); Ur, *Umbelopsis ramanniana*.

		BnDGAT2a	BnDGAT2b
Plant	AtDGAT2	76.3%	75.7%
	RcDGAT2	54.7%	54.4%
	VfDGAT2	57.8%	57.8%
Animal	HsDGAT2	22.7%	23.2%
	MmDGAT2	20.9%	21.6%
Fungus	ScDGAT2	20.1%	20.3%
	SpDGAT2	24.2%	26.7%
	UrDGAT2a	23.1%	23.7%
	UrDGAT2b	25.1%	26.5%

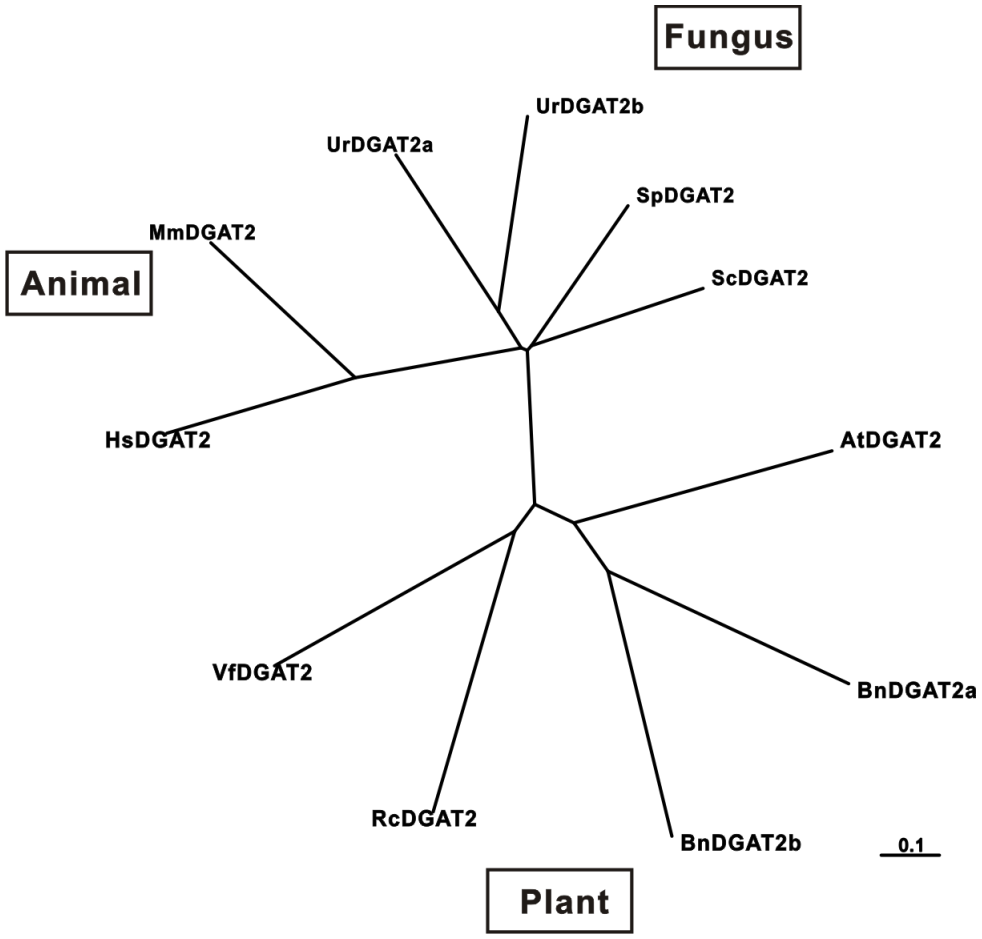


Figure 3.1. Phylogenetic analysis of the acyl-CoA:diacylglycerol acyltransferase (DGAT2) family. The branch lengths of the tree are proportional to divergence. The 0.1 scale represents 10% change. Proteins used in the analysis were DGAT2 sequences from human (*Homo sapiens*), Hs; mice (*Mus musculus*), Mm; *A. thaliana*, At; *B. napus*, Bn; tung (*V. fordii*),Vf; castor bean (*R. communis*), Rc; budding yeast (*S. cerevisiae*), Sc; fission yeast (*Schizosaccharomyces pombe*), Sp; *Umbelopsis ramanniana*, Ur, *Umbelopsis ramanniana*.

3.3.2 Expression of *DGAT2* candidate genes in yeast

To select a functional DGAT2 candidate protein, various *DGAT2* cDNAs were obtained from available sources including plant (*A.thaliana* and *B.napus*) as well as yeast (*S.cerevisiae*) followed by expression in yeast strain H1246 as described under "Materials and methods." Total RNA was extracted from yeast cells expressing *AtDGAT2*, *BnDGAT2*s and *ScDGAT2*, respectively. RT-PCR was performed using primers specific for each gene to investigate the transcriptional expression of candidate genes. *Actin*, as a housekeeping gene, was used as a control.

As shown in Figure 3.2A, transcripts of all genes encoding AtDGAT2, BnDGAT2s and ScDGAT2 could be detected at significant levels. Expression at the translational level was monitored by immunoblotting using anti-V5 antibodies which recognized the engineered C-terminal V5-epitope fused to the recombinant proteins. Identity of ScDGAT2 expression was verified by immunoblotting detection of a single band in the microsomal fraction with an approximate molecular weight of 48kDa (Figure 3.2B, lane 4 and 5). The polypeptides corresponding to plant DGAT2s, however, could not be observed in either microsome or supernatant fractions (Figure 3.2B, lane 6-11).

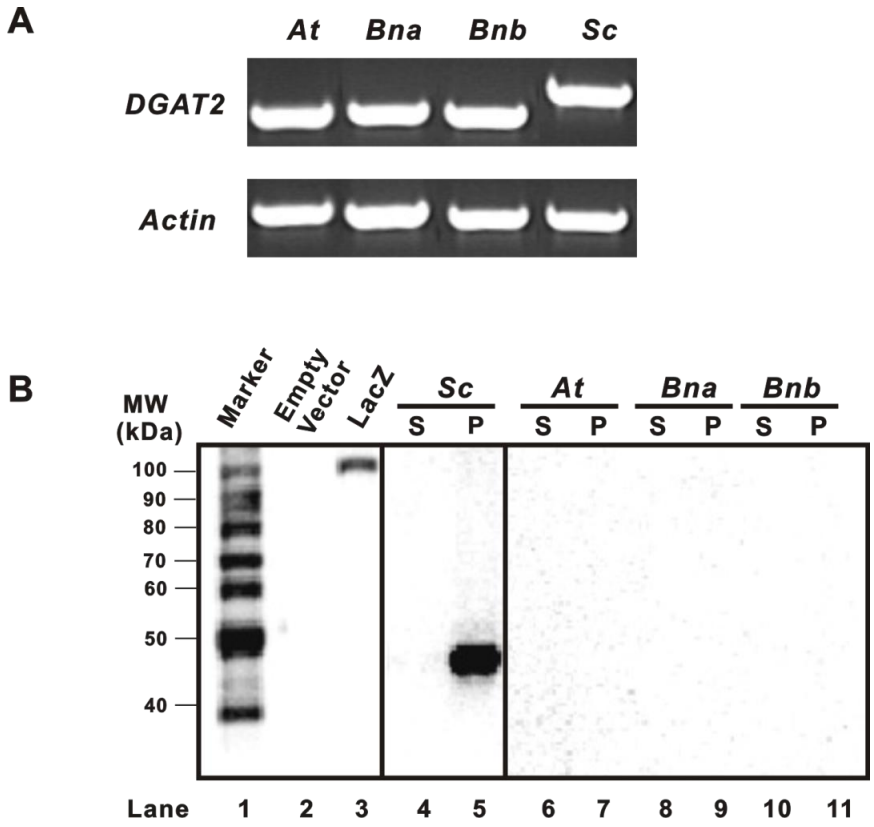


Figure 3.2. Expression of different *DGAT2*s in yeast H1246 cells. (A) RT-PCR analysis of gene expression. *Bna* and *Bnb* represent the two isoforms of full-length *BnDGAT2* cDNA which were deposited in the GenBank under the accession numbers FJ858270 and FJ858271, respectively. (B) Immunoblot of DGAT2s probed with anti-V5 antibodies. Membranes were isolated from the yeast strain H1246 cells expressing V5-tagged ScDGAT2, AtDGAT2 and BnDGAT2s. Microsome pellet (P) was separated from cytosolic protein in supernantant (S) by 100,000×*g* centrifugation. Equal amounts of protein from each fraction were resolved by 10% SDS-PAGE and analyzed by immunoblotting.

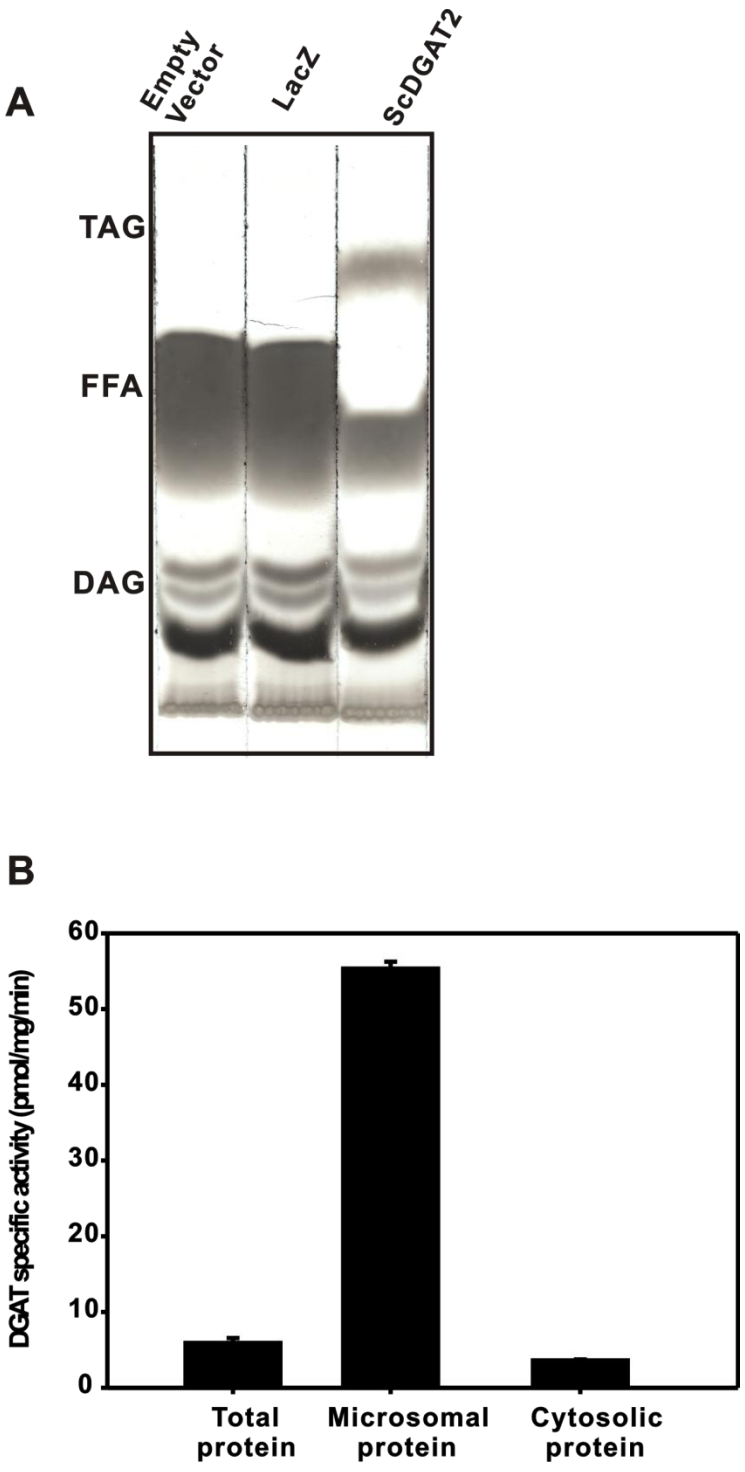
3.3.3 Functional analysis of DGAT2 protein candidates

The biological activities of the recombinant proteins were probed by performing lipid analysis and DGAT *in vitro* enzyme assay. Total lipid was extracted from H1246 yeast cells expressing different *DGAT2* cDNAs and separated by TLC. The yeast strain H1246, devoid of endogenous neutral lipid synthesis, has been extensively used as a tool to study TAG metabolism (Table 2.1). As shown in Figure 3.3A, H1246 yeast cells transformed with control plasmids including empty vector or *pYESLacZ* did not produce TAG, as expected in the quadruple knock-out background. Only the cells transformed with plasmid-borne *ScDGAT2* could restore TAG biosynthesis *in vivo*. Overexpression of *AtDGAT2* and *BnDGAT2*s in H1246 cells, however, did not lead to a detectable TAG band (data not shown).

An *in vitro* enzyme assay was also performed to investigate the activities of DGAT2 candidate proteins. Neither soluble nor microsomal protein preparations from yeast cells expressing plant *DGAT2* cDNAs showed detectable DGAT2 activity in our system (data not shown). DGAT activity in yeast cells transformed with empty vector or *pYESLacZ* was also not observed, as expected. In contrast, substantial ScDGAT2 activity was present predominantly in the microsomal fraction at a specific activity of 54 pmol of TAG formed per mg of

protein/min, which was five times higher than that found in cytosolic and total protein fractions (Figure 3.3B).

Figure 3.3. Functional analysis of recombinant ScDGAT2. (A) TLC separation of different lipid classes in total lipids extracted from yeast cells. Position of lipid standards are indicated at left as follows: diacylglycerol (DAG), free fatty acid (FFA) and triacylglycerol (TAG). Results are representative of two independent experiments. (B) *In vitro* DGAT activities in lysate (total protein), cytosolic (cytsosolic protein) and microsomal fraction (microsomal protein) from yeast cells expressing *ScDGAT2*. Different protein fractions from yeast cells expressing *DGAT2* were incubated in the presence of *sn*-1,2-dioleoylglycerol and [1-^{14}C] oleoyl-CoA and assayed as detailed under "Materials and methods." All enzyme activities were expressed as picomoles of TAG formed per minute per milligram of yeast microsomal protein as a mean± S.D for biological triplicate analysis in one experiment.


A
Empty Vector
LacZ
ScDGAT2
TAG
FFA
DAG
B
DGAT specific activity (pmol/mg/min)
60
50
40
30
20
10
0
Total protein
Microsomal protein
Cytosolic protein

3.3.4 Optimization of *ScDGAT2* expression

The previous results indicated that *ScDGAT2* was the only cDNA which resulted in the production of detectable and functional enzyme. Therefore, ScDGAT was chosen for further functional and structural studies. To optimize *ScDGAT2* expression, a time-course expression analysis was performed wherein enzyme production levels were monitored by enzyme assay combined with immunoblotting analysis. As demonstrated in Figure 3.4, optimal levels of polypeptide production were attained at 18 hours post-induction with a good correlation between the protein accumulation and the level of DGAT2 activity over the same time period. The effect of C-terminal tag on enzyme function was examined by constructing an untagged ScDGAT2. Assays on both versions of DGAT2 showed that V5-tagged DGAT2 was about 40% less active than the untagged enzyme (Figure 3.5).

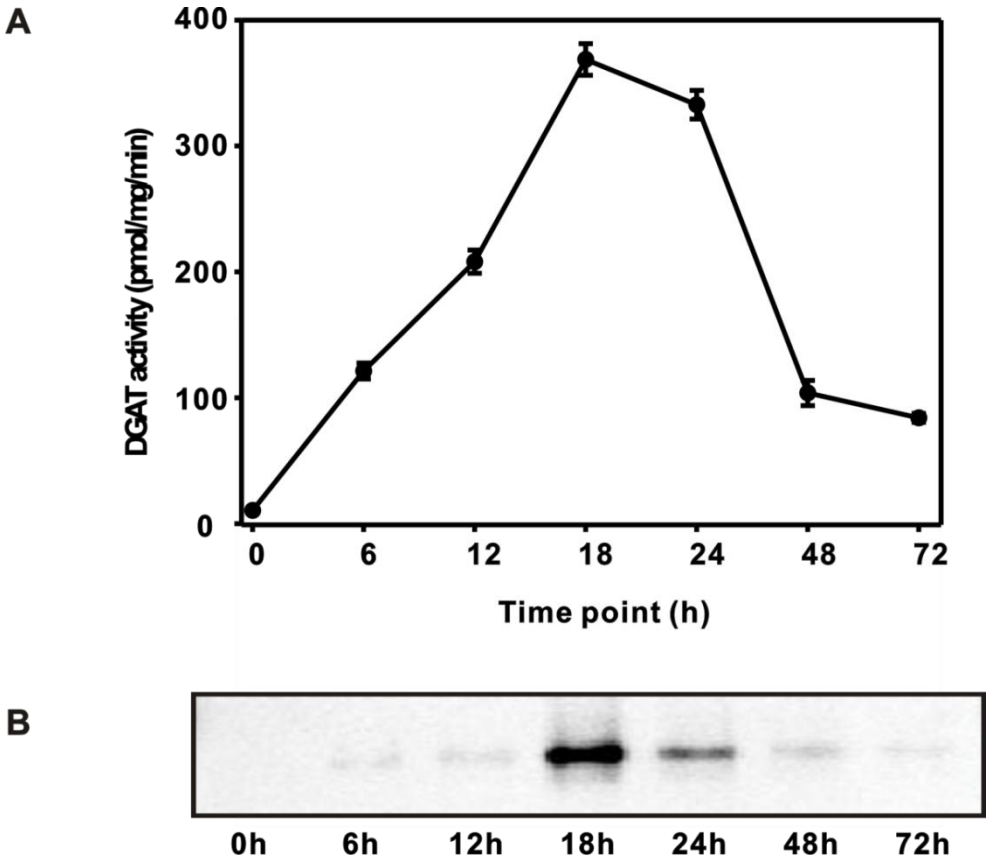


Figure 3.4. Time course of the polypeptide expression of ScDGAT2. (A) DGAT activities in microsomes from yeast cells expressing V5-tagged ScDGAT2, harvested at various times after induction of expression. The *in vitro* enzyme assay was performed under the same conditions as for the experiments described in Figure 3.3B. Data represent the average of three assays in one experiment, with the error bars corresponding to the S.D. (B) Immunoblots of microsomal protein of the same batch of cells harvested at the same times used for enzyme assays.

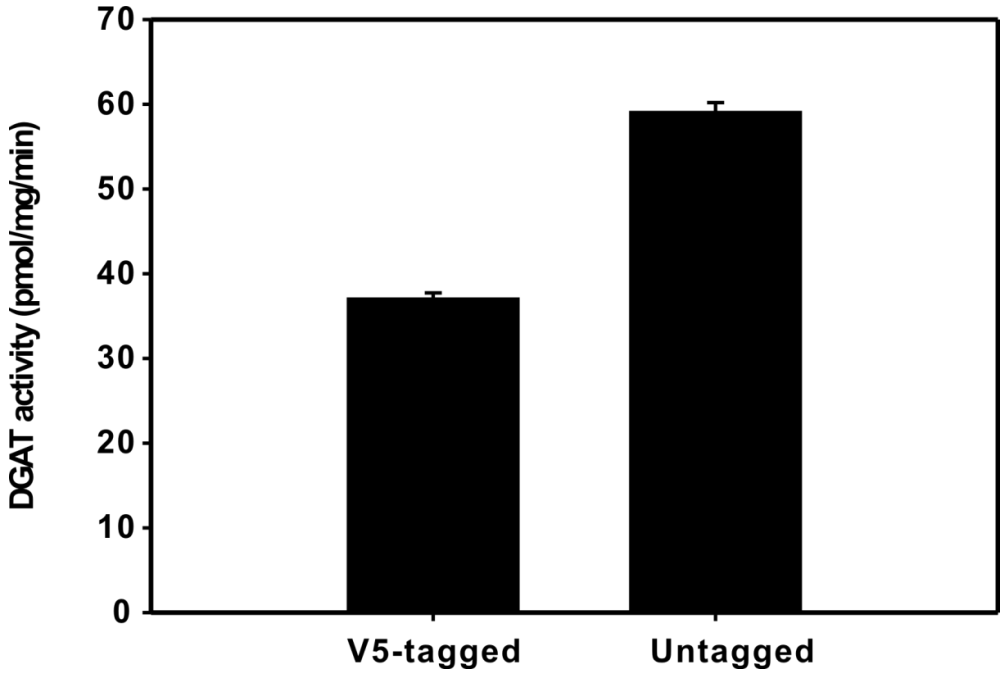


Figure 3.5. Activity of untagged ScDGAT2 compared to V5-tagged ScDGAT2. Untagged or V5-tagged ScDGAT2 constructs were prepared as described under "Materials and methods." Microsomal membranes were isolated from the yeast strain H1246 expressing either version of *ScDGAT2* followed by enzyme assay. The conditions for *in vitro* enzyme assay are the same as for the experiments described in Figure 3.3B. Data represent means ± S.D. (n=3).

3.4 Discussion

The biochemical activity of DGAT associated with TAG bioassembly is well studied, but few insights have been reported on molecular mechanisms of this class of enzyme due to absence of high-resolution structural data and technical challenges associated with characterizing membrane proteins. To begin addressing this issue, we proposed using a combination of site-directed mutagenesis (SDM) and chemical modification to investigate the structural and functional relationships of DGAT2, which, at the time, was believed to have a simpler structure compared to DGAT1 (discussed in Chapter 2). As an initial step, it was necessary to select a *DGAT2* cDNA that could be functionally expressed in a recombinant system.

Although DGAT2 has been cloned and functionally characterized in mammals and fungi for many years, no functional orthologues had been identified in plants when this project began. Thus, we initially attempted to express a cDNA encoding putative DGAT2s from the model plant *A. thaliana* and the oilseed crop *B. napus*, the latter in which we have been studying TAG biosynthesis in our laboratory (Lung and Weselake, 2006; Snyder et al., 2009; Weselake et al., 2000). Based on DNA homology to other available plant *DGAT2* sequences, two isoforms of *DGAT2* cDNA encoding polypeptides with highly identical deduced

amino acid sequences were isolated from *B. napus*. These deduced amino acid sequences were substantially homologous with other plant DGAT2s.

AtDGAT2 and *BnDGAT2*s were successfully transcribed in yeast cells, but their expression was blocked at the translational level, resulting in no protein accumulation. Since these plant DGAT2s were heterologously expressed in yeast cells, there may be other unidentified factors existing only in these plant species which are needed to maintain the protein stability. It could be also the different preference in codon usage between the yeast and plant organisms (Angov et al., 2008; Chiapello et al., 1998; Sharp et al., 1988; Sharp et al., 1986). This lack of protein production is probably the reason why both plant DGAT2s showed no detectable activity in *in vitro* enzyme assay and no complementation effect on TAG formation in yeast cells. These findings about expression of plant *DGAT2s* are in line with recent work *in A. thaliana* (Li et al., 2010; Zhang et al., 2009). Transcriptional expression of *AtDGAT2* was detected in many tissues including developing seeds, although *AtDGAT2* was expressed at lower level compared to *AtDGAT1* (Li et al., 2010). A T-DNA insertion mutant of *AtDGAT2*, however, did not show any reduction in the TAG content compared to the wild type *A. thaliana* (Li et al., 2010; Zhang et al., 2009). In addition, double mutant of *AtDGAT1* and *AtDGAT2* did not have a substantial additional decrease in oil content compared with the *DGAT1* single mutant (Zhang et al., 2009). Nevertheless, we cannot

exclude the possiblity that knockout of *DGAT* by mutagenesis could be partially compensated by other genes related to TAG metabolism. Thus, DGAT2 might be involved with TAG metabolism at some level but it does not play as a substantial role as DGAT1 in certain plant species. More advanced studies at both genetic and biochemical levels in plants, such as mutiple gene disruptions and protein-protein interactions, will help understand the relative contribution of DGAT2 in the TAG assembly.

Plant *DGAT2*s tested in this study could not be expressed at translational levels in yeast, but *ScDGAT2* was revealed to be functionally expressed and accumulated in the microsomal fraction of *S. cerevisiae*. These findings are consistent with previous reports on other DGAT2 family members such as DGAT2 from *Umbelopsis ramanniana*, castor bean (*Ricinus communis*) and tung tree (*Vernicia fordii*), supporting the hypothesis that DGAT2 is associated with the ER (Kroon et al., 2006; Lardizabal et al., 2001; Shockey et al., 2006). The low amount of DGAT activity in the cytosol of yeast H1246 cells expressing *ScDGAT2*, as shown in Figure 3.3B, suggests that some unsedimented membrane fragments may still remain in the cytosol and/or that the recombinant enzyme may also be associated with lipid bodies which would not have sedimented during ultracentrifugation. Indeed, a previous study demonstrated that the yeast DGAT2 is also active in the lipid droplet (Sorger and Daum, 2002). Although expression

of *ScDGAT2* in H1246 resulted in formation of endogenous TAG, lipid droplets were not formed to a level that allowed their isolation.

In conclusion, the plant *DGAT2* cDNAs examined (from *A.thaliana* and *B. napus*) did not result in the production of detectable polypeptide or enzyme activity when the cDNAs were expressed in H1246 *S. cerevisiae*. In contrast, cDNA encoding ScDGAT2 was functionally expressed in yeast and the resulting polypeptide accumulated in the microsomal fraction. Thus, this latter system was chosen for structure/function studies wherein putative functional motifs/residues are probed using a combination of SDM and chemical modification. The results of these experiments are described in subsequent chapters.

A version of this chapter has been published. Liu, Q., Siloto R. M. P. and Weselake R. J. 2010.

Biochemistry. 49: 3237–3245

4. Probing the Role of Cysteine Residues in Thiol-modification on Acyl-CoA:Diacylglycerol Acyltransferase 2

4.1 Introduction

Triacylglycerol (TAG or TG) is a neutral lipid which primarily provides reservoirs for energy storage and membrane formation in most of eukaryotic organisms. Disorders of TAG metabolism in different human tissues are associated with a variety of diseases, such as obesity, type-II diabetes, non-alcoholic fatty liver and coronary heart disease (Millar and Billheimer, 2006; Mulhall et al., 2002; Stone et al., 2004). A better understanding of TAG biosynthesis and the roles of the enzymes involved in this metabolic pathway could hasten the development of therapeutic intervention for treating these disorders and may aid in the growth of new biotechnological tools for production of value-added oils in plants and micro-organism.

Acyl-CoA:diacylglycerol acyltransferase (DGAT or DAGAT, EC 2.3.1.20) is an enzyme mainly associated with the endoplasmic reticulum (ER) which catalyzes the terminal step in TAG biosynthesis (Lung and Weselake, 2006; Siloto et al., 2009a; Yen et al., 2008). Several lines of evidence suggest that the level of DGAT activity may have a substantial effect on carbon flow into

TAG, indicating that DGAT is critical for regulating TAG metabolism (Bagnato and Igal, 2003; Cahoon et al., 2007; Millar et al., 2006). There are at least two distinct DGAT families, DGAT1 and DGAT2, with gene orthologues in animals, plants and microorganisms (Yen et al., 2008). DGAT1 and DGAT2 have strikingly different physiological roles. Reverse genetics studies have shown that DGAT2 knockout mice are severely deficient in TAG and manifest a skin barrier dysfunction, leading to early death. In contrast, DGAT1-deficient mice are viable although they have reduced TAG accumulation in tissues (Smith et al., 2000; Stone et al., 2004). DGAT1 is about 500 amino acid residues in length and has been proved to have only three transmembrane domains (TMDs) (McFie et al., 2010). DGAT2 has about 400 amino acid residues, is less hydrophobic and may have only one or two TMDs (Stone et al., 2006). Further insights into the molecular mechanism and the assignment of functional roles to structural domains of DGATs, however, have been limited because of the complexities in the solubilization and purification of these membrane-bound enzymes (Little et al., 1994). To date, the high-resolution structure of either DGAT member remains unknown.

Current understanding of DGAT function has been restricted to local functional information obtained by biochemical approaches using microsomal fractions (summarized in Chapter 2). *In vitro* chemical modification studies have

been employed for characterizing DGAT for many years. Sauro and Strickland (1990) have shown that DGAT activity of rat myotube tissue is sensitive to thiol-specific reagents such as N-ethylmaleimide (NEM) and p-chloromercuribenzoate. DGAT activity from the fungus *Umbelopsis ramanniana* (formerly *Mortierella*) was also inhibited by NEM (Kamisaka et al., 1993). Our previous results on microsomal DGAT activity from muscle tissue of *Bos taurus* were in agreement with these findings (Lozeman et al., 2001). Since microsomal fractions are likely to contain both DGAT1 and DGAT2 polypeptides, it is difficult to ascribe the effects of inhibition to one specific DGAT isoform (Lung and Weselake, 2006; Yen et al., 2008). Nevertheless, these data suggest that cysteine residues, which are frequently present in functionally important sites because of their high reactivity, may play a functional role in DGAT1 and/or DGAT2. Therefore, cysteine residues may represent valuable targets for investigating structure and function in these enzymes.

In this work, we aimed to probe the functional and structural roles of cysteine residues in DGAT2. Based on our work detailed in chapter 3, we chose the DGAT2 orthologue from yeast (*Saccharomyce cerevisiae*), also known as DAGAT or Dga1p (referred to here as ScDGAT2) as a model protein. This enzyme is the only DGAT characterized in *S. cerevisiae* and is responsible for most of the TAG production in this organism (Oelkers et al., 2002; Sorger and

Daum, 2002). We demonstrated that ScDGAT2 can be inactivated by different thiol-modifying reagents. We then undertook a detailed, systematic investigation of the mechanisms of thiol-modification on yeast DGAT2 by probing the structural and functional role of seven cysteine residues in this enzyme (Cys^{48}, Cys^{49}, Cys^{127}, Cys^{183}, Cys^{211}, Cys^{264} and Cys^{314}) using cysteine-specific modification and site-directed mutagenesis (SDM). Although none of the cysteines appear to be directly involved in DGAT activity, our results located a putative functional motif in DGAT2.

4.2 Materials and methods

4.2.1 Materials

Iodoacetamide (IA), NEM, SDS and triolein were from Sigma. 5, 5'-dithiobis-2-nitrobenzoate (DTNB) was from Pierce. mPEG5000-maleimide (PEG-mal) was purchased from Sunbio (Orinda, CA). [1-^{14}C]Oleoyl-CoA was obtained from GE Healthcare (Baie d'Urfe, Quebec, Canada). Diolein was from Avanti (Alabaster, AL). Primers were synthesized by IDT (Coralville, IA). Anti-V5-HRP antibodies were from Invitrogen. PVDF membrane (hybond-P) and the enhanced chemiluminescence kit (ECL advance) were obtained from GE Healthcare. Yeast media and chemicals were purchased from BD Difco (Oakville, Ontario, Canada) or Fisher Scientific (Whitby, Ontario, Canada).

4.2.2 Strain and cell culturing

The *S. cerevisiae* stain with quadruple knockout of the *DGA1*, *LRO1*, *ARE1* and *ARE2* (H1246, *MATα are1*-Δ*::HIS3, are2*-Δ*::LEU2, dga1*-Δ*::KanMX4, lro1*-Δ*::TRP1 ADE2*) was used as the host strain for protein expression (Sandager et al., 2002). Yeast transformation was performed according to Gietz and Schiestl (2007). Single colonies were inoculated into minimal media containing 0.67% (w/v) yeast nitrogen base, 2% (w/v) dextrose, 20mg/L of adenine, arginine, tryptophan, methionine, histidine, and tyrosine, 30mg/L of lysine and 100mg/L of leucine. Yeast cells in liquid media were cultured at 30^{o}C and 250 rpm in an orbital shaker overnight. An appropriate volume of cell culture was harvested and resuspended in induction medium containing 2% (w/v) galactose and 1% (w/v) raffinose instead of dextrose at OD600 of 0.4 to induce the expression of recombinant genes.

4.2.3 Preparation of ScDGAT2 mutants

SDM within the ScDGAT2 ORF was generated using PCR-driven overlapping extension described elsewhere (Heckman and Pease, 2007). Briefly, high fidelity *platinum* Taq DNA polymerase (Invitrogen, Carlsbad, CA) was employed to amplify two overlapping PCR fragments. A list of ScDGAT2 mutants with different amino acid substitutions and corresponding primers used for mutagenesis are described in Table 4.1. The full coding region of each mutant

was then reconstructed using these two fragments as templates. The mutant C0, in which all 7 cysteines were converted to alanines, was synthesized by IDT and was used as the template for construction of mutant A314C. Other mutants were prepared using the template provided by ScDGAT2 in pYES2.1-TOPO. All mutations were confirmed by DNA sequencing.

Table 4.1. Mutagenic oligonucleotides used for site-directed mutations within ScDGAT2

Mutation	Forward	Reverse
CC48-49AA	CTCAAACCACAACTAGAGTCAGCAGCTCCATTGGCGACC	GGTCGCCAATGGAGCTGCTGACTCTAGTTGTGGTTTGAG
C127A	CATTTGGAAGTGGTATGCTGATTATTTC	GAAATAATCAGCATACCACTTCCAAATG
C183A	TGACTATCGCAACCAGGAAGCTACAGGGCCAAC	GTTGGCCCTGTAGCTTCCTGGTTGCGATAGTC
C211S	GAGCGTTTGCAACAGAAGGTTCTAACTATTCCAAG	CTTGGAATAGTTAGAACCTTCTGTTGCAAACGCTC
C264A	TAAGCAAAAATCAGTCGATCGCCATTGTTGTTG	CAACAACAATGGCGATCGACTGATTTTTGC
C314A	TTGCATTTGGAGAGGTGGACGCTTATAATGTTC	GAACATTATAAGCGTCCACCTCTCCAAATGC
*A314C	TTGGAGAGGTGGACTGTTATAATGTTCTGAGC	GCTCAGAACATTATAACAGTCCACCTCTCCAA

Modification was indicated by an underline and sequence are written from 5’ to 3’

* This mutation was prepared using a Cys-less mutant of ScDGAT2 as the template

4.2.4 DGAT assay

Microsomal proteins were prepared from induced yeast H1246 cells by glass beads as detailed in Chapter 3. Concentration of microsomal protein was determined by BCA assay (Pierce, Rockford, IL). *In vitro* DGAT assays were performed according to the method of Weselake et al. (1991) with the following modifications. A final volume of 55μL assay mixture contained 10μg of the microsomal protein, 0.15M Hepes-NaOH (pH7.0), 2.5mM $MgCl_2$, 0.1mg/mL BSA (fatty acid free), 308μM *sn*-1,2-diolein and 15μM [1-^{14}C] oleoyl-CoA. Assay mixtures were incubated at 30ºC for 10min, and the reactions were terminated by adding10μL 10% SDS (w/v). Fifty microliters of the terminated reaction mixture were applied to silica gel TLC plates. The TLC plates were developed in n-hexane:diethyl ether:acetic acid (80:20:1, v/v) and TAG spots were visualized with iodine. Confirmation of radiolabeled TAG identity was determined by comigration of unlabeled triolein carrier. Radioactivity was quantified by a Beckman-Coulter LS6500 liquid scintillation counter. Assays were performed in three biological replicates. For enzyme inhibition assays, thiol protection and modification reagents dithiothreitol (DTT), IA, NEM and DTNB were prepared individually as stock solutions in dimethyl sulfoxide (DMSO). For each reaction, 50μL of microsomes with concentration of 1mg/mL were pre-incubated with 1μL of a thiol-modifying reagent at 4^{o}C for 30min prior to

initiation of the assay. Enzyme activity based on incubation with 1μL DMSO served as the control with values of 100%.

4.2.5 Thiol-specific chemical modification

Direct and indirect thiol-specific chemical modification with PEG-mal and IA under denaturing condition was performed as reported previously (Guo et al., 2005a; Lu and Deutsch, 2001) with the following modifications. Stock solutions of thiol-modifying reagents IA and PEG-mal were fresh prepared in the extraction buffer (0.2M Hepes-NaOH, 0.5M sucrose, 1mM phenylmethylsulphonyl fluoride, pH 8.7) with 10% (w/v) SDS. For the direct modification, 150μg of cell homogenate were simultaneously added with different concentrations of IA and PEG as indicated in a final volume of 100μL and incubated at 37^{o}C for 30min. Under these conditions, ScDGAT2 was fully denatured in 5% SDS. Free cysteines were efficiently modified by either IA or PEG-mal and an appropriate amount of incubation mixture was used for immunoblotting. For the indirect modification, 100μL of 20mM IA in buffer with 10% SDS was first added to the same volume of cell homogenate (3mg/mL) at 37^{o}C for 30min to block the free cysteines in the protein. Excess IA was then removed by trichloroacetic acid precipitation. Recovered pellets were resuspended in 50μL extraction buffer (containing 5% SDS) with DTT (8mM) or 2-mercaptoethnol (2-ME, 10mM) and a sample without reducing reagents was used as a negative control. After

incubation at 37^{o}C for 30min, 50μL of 8mM PEG-mal were added to modify the free cysteines which would have been involved in disulfide bond before reduction. Following additional 30min incubation, the reaction mixture was then analyzed by SDS-PAGE and immunoblotting. The modification of free cysteines in these processes was monitored by quantifying the total free sulfhydryl groups using DTNB reagent. Cysteine hydrochloride monohydrate was used as the standard for quantification.

4.2.6 SDS-PAGE and immunoblotting analysis

Protein samples were mixed with loading buffer (125mM Tris, pH6.8, 2% SDS, 25% glycerol, 5% 2-mercaptoethanol and 0.15% bromophenol blue), incubated at room temperature for 30min and analyzed by SDS-PAGE (10%). For samples used for mapping disulfide bonds, an appropriate amount of the modified homogenate was added to the loading buffer without 2-ME. Immunoblots were carried out by transferring proteins to a PVDF membrane. The recombinant ScDGAT2 was detected using anti-V5-HRP antibodies (Invitrogen) according to the protocol described by the manufacturer. Densitometry analysis of immunoblots was performed using ImageQuant TL software (Amersham).

4.3 Results

4.3.1 ScDGAT2 activity is susceptible to thiol-modifying reagents

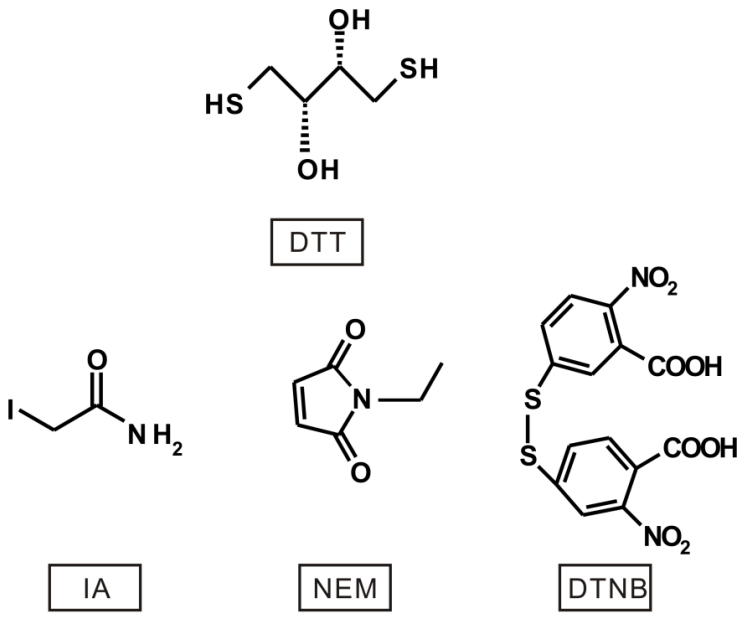
The scope of this study was to understand the role of cysteines in the structure and function of a DGAT2. It was important to ensure that the enzyme is functionally active. As described in Chapter 3, we have revealed that ScDGAT2 was functionally expressed and accumulated in the microsomal fraction of *S. cerevisiae*. Thus, we used the microsomal vesicles as the source of recombinant ScDGAT2 in this work.

To investigate the influence of cysteines in the enzyme activity of ScDGAT2, we evaluated the effects of thiol-protecting reagent DTT and thiol-modifying reagents IA, NEM and DTNB which have different structures and properties (Figure 4.1A). Treatments of ScDGAT2 in yeast microsomes with DTT resulted in a marginal increase of enzyme specific activity because it could prevent formation of disulfide bond between free cysteines in the protein as a reducing reagent. IA, NEM and DTNB (commonly known as Ellman's reagent) represent three different classes of cysteine modifiers which are frequently used in protein biochemistry: halogen-beta-ketone, maleimide and aromatic disulfide. Modification with the NEM (membrane-permeant) and DTNB (membrane-impermeant) decreased ScDGAT2 activity in a concentration-dependent manner with the IC_{50} values of 60 μM and 35 μM, respectively. NEM modifies the thiols

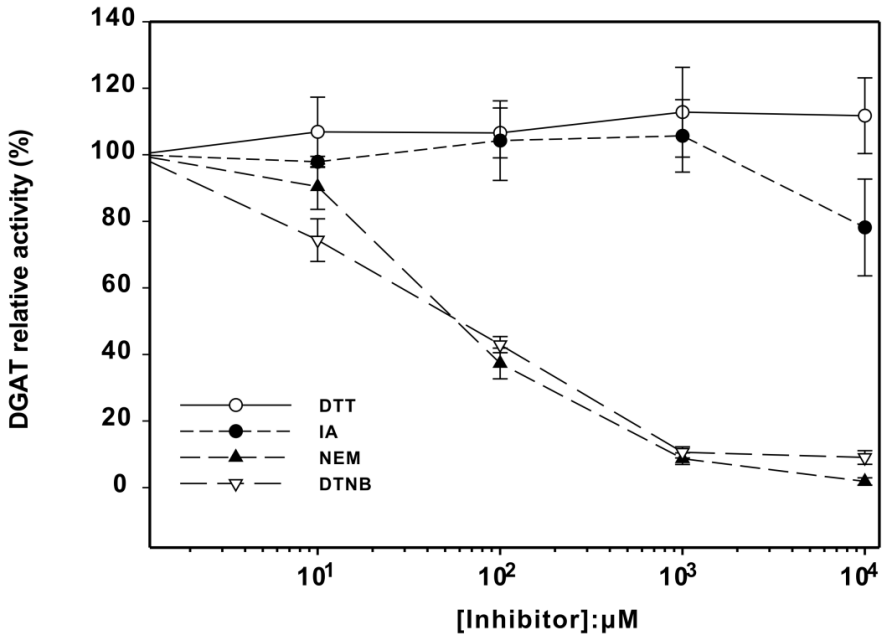
by alkylation while DTNB acts by cleaving its internal disulfide bond. IA, which is also membrane-permeant and reacts with free cysteines in the same way as NEM, inhibited the enzyme activity only at concentrations above 1mM (Figure 4.1B). The main difference between IA and other two modifiers is that IA has a relatively small functional group. These results raised the hypothesis that certain reduced cysteines might be involved in the activity of ScDGAT2.

Figure 4.1. Susceptibility of ScDGAT2 enzyme activity to thiol-modifying reagents. (A) Chemical structures of thiol-specific reagent dithiothreitol (DTT), iodoacetamide (IA), N-ethylmaleimide (NEM) and 5, 5'-dithiobis-2-nitrobenzoate (DTNB). (B) Microsome fractions containing V5-tagged ScDGAT2 were treated with DTT, IA, NEM or DTNB followed by DGAT assay. Results are expressed as percent of residual activity, i.e. percent of the activity in the presence of modification reagent with respect to the control value (set as 100%) in which no thiol-specific reagents were added. Values of reagent concentration were log-normalized to evaluate its correlation to residual activity. Three independent treatments were assayed in one experiment. Error bars indicate the S.D of the mean. A four-parameter logistic model was used to fit to the data to calculate the IC_{50} values (http://bsmdb.tmd.ac.jp:3000/cbdb/ic50).

A


OH
SH
HS
OH
DTT
O
I
NH2
IA
O
N
O
NEM
NO2
COOH
S
S
COOH
NO2
DTNB

B

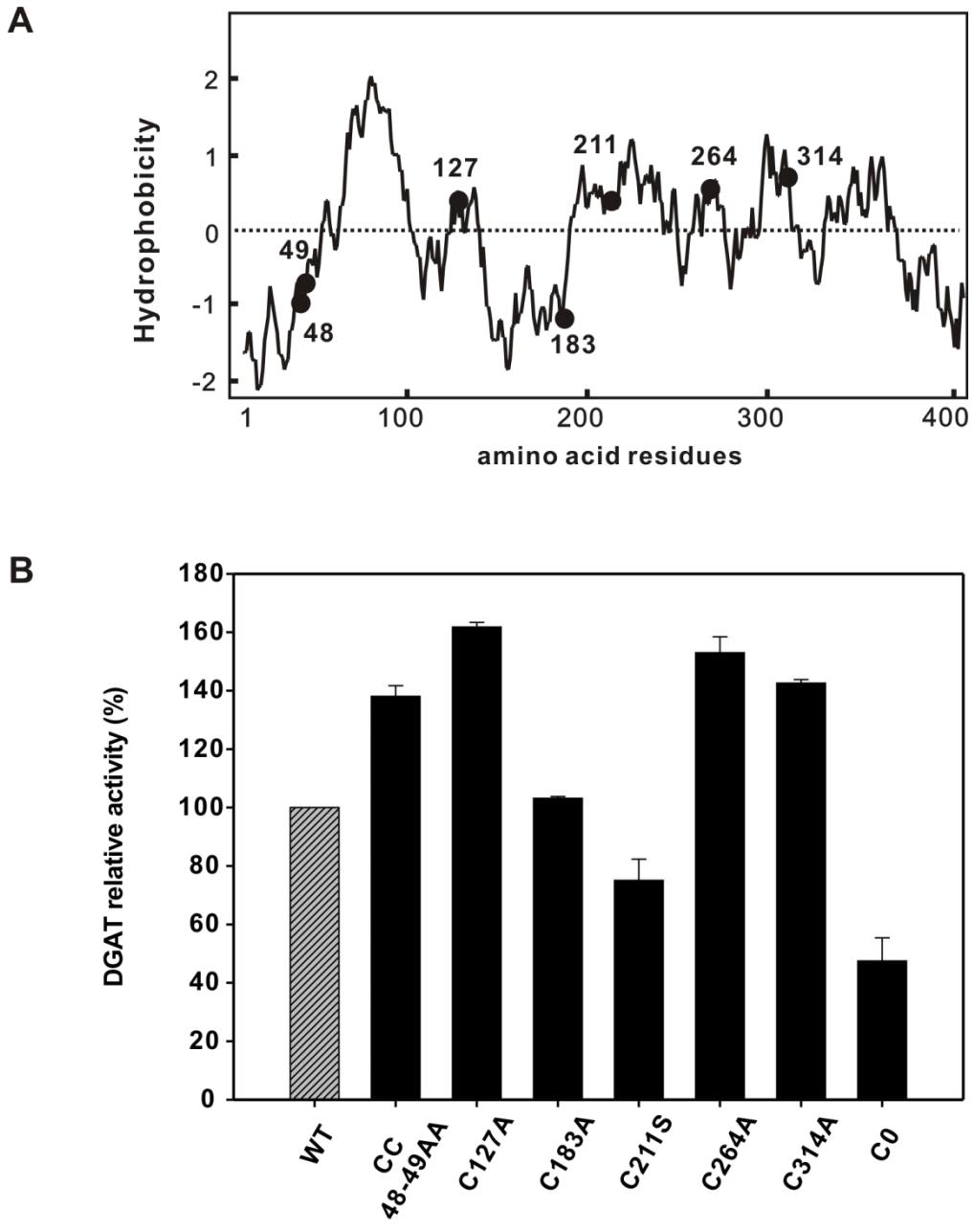

DGAT relative activity (%)
140
120
100
80
60
40
20
0
DTT
IA
NEM
DTNB
10^1
10^2
10^3
10^4
[Inhibitor]:μM

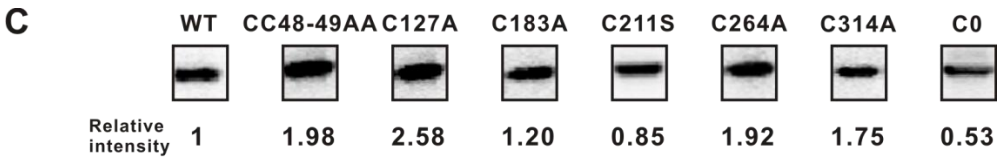
4.3.2 Cysteine residues are not essential for enzyme catalysis

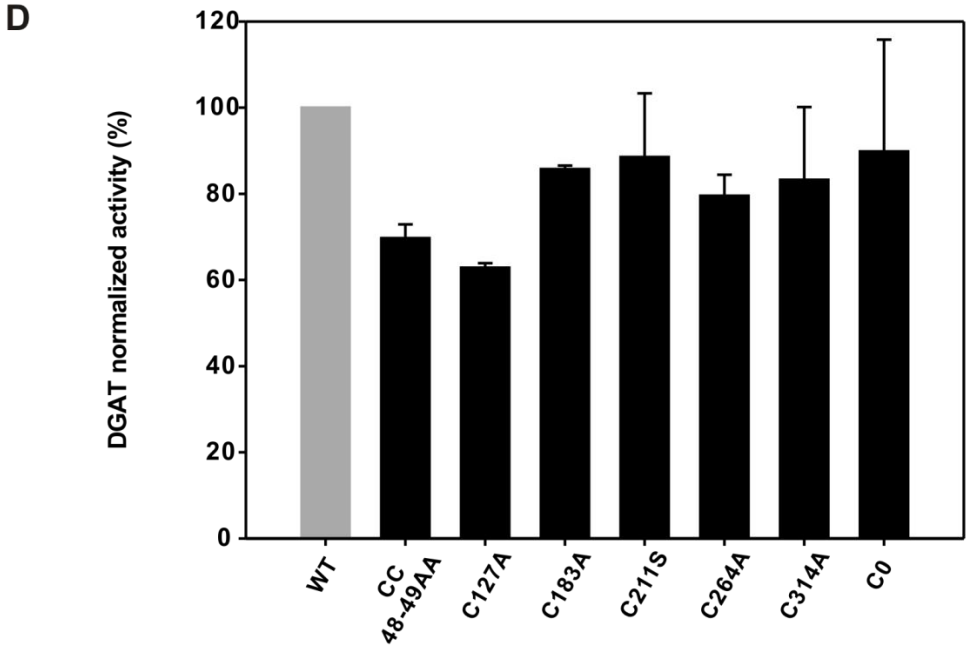
In order to further explore the role of each of the seven cysteines in ScDGAT2 activity, we examined the hydrophobicity of the position corresponding to each cysteine residue by generating a Kyte-Doolittle (1982) hydropathy plot (Figure 4.2A). We then applied an SDM approach to generate the mutants C127A, C183A, C264A and C314A. The two contiguous cysteines (Cys^{48} and Cys^{49}), were also replaced with alanines to produce the mutant CC48-49AA. Cys^{211} is situated in a hydrophobic segment (Figure 4.2A); therefore, to minimize disruption to protein conformation, we converted Cys^{211} to serine replacing the thiol with a hydroxyl group, yielding the mutant C211S. Each ScDGAT2 mutant was expressed in the strain H1246 as previously described. The specific activity for each mutant was measured and the results suggested that all mutants displayed equal or higher specific activity than wild type with the exception of C211S (Figure 4.2B). Immunoblots showed, however, that each recombinant protein accumulated at a different level as determined by densitometry analysis where the intensity of wild type ScDGAT2 was set as 1 (Figure 4.2C). After normalizing the specific activities with the immunobloting analysis, we found that the relative activity of each mutant was comparable to wild type except for CC48-49AA and C127A, which have mild reduction in enzyme activity (Figure 4.2D). Therefore, single cysteine to alanine replacements

might only have induced subtle alterations in the protein architecture. To investigate the possibility of a synergistic action of different cysteines, we also generated the mutant C0, where all cysteines were converted to alanines. C0 was expressed in the strain H1246 and analyzed as described for the previous mutants. Although C0 accumulated at lower levels compared to wild type, it was functionally active (Figure 4.2C and 4.2D), demonstrating that cysteines are not essential for DGAT2 catalysis. To validate the role of the cysteine residues in the inactivation of ScDGAT2, we repeated the treatment with the same thiol-modifying reagents using the functional mutant C0. As shown in Figure 4.3, DTT, IA or NEM did not produce any significant effect on mutant C0 activity. DTNB, however, decreased enzyme activity in concentration-dependent manner, with over 30% inhibition at a reagent concentration of 1mM. These results confirmed that the decrease in enzyme activity observed with IA and NEM treatments were attributable to the modification of cysteine residues. The inhibitory effects of DTNB on the C0 mutant, however, suggested that this reagent has an unrelated nonspecific effect on enzyme activity.

Figure 4.2. Expression of cysteine to alanine mutants. (A) Hydropathy plot of ScDGAT2 showing different hydrophobicity environments of each cysteine residue in ScDGAT2. The plot was generated by the method of Kyte and Doolittle (1982) using a window size of 19. Cysteine residues in ScDGAT2 are represented by circles and the positions were numbered. (B) DGAT relative activities of microsome isolated from cells expressing cysteine mutant CC48-49AA, C127A, C183A, C211S, C264A, C314A and Cys-less mutant C0. Details for designing these mutants are described under "Results." All activities are quoted relative to that of the wild type (WT) protein (defined as 100%). Data represent means ± S.D of experiments in triplicate. (C) Expression level of each mutant was examined by immunoblotting using the same amounts of protein from the same batch of microsomal proteins as used for DGAT assay. The results shown are from one of three representative experiments. The relative band intensities compared to that of wild type (set as 1) estimated by ImageQuant TL (GE Healthcare) are shown as the mean of triplicate. (D) Normalized activity of ScDGAT2 cysteine mutants. DGAT relative activity was normalized by dividing the activity value with the relative intensity.


A
Hydrophobicity
2
1
0
-1
-2
49
48
127
211
183
264
314
1
100
200
300
400
amino acid residues
B
DGAT relative activity (%)
180
160
140
120
100
80
60
40
20
0
WT
CC 48-49AA
C127A
C183A
C211S
C264A
C314A
C0


C
WT
CC48-49AA
C127A
C183A
C211S
C264A
C314A
C0
Relative intensity
1
1.98
2.58
1.20
0.85
1.92
1.75
0.53


D
DGAT normalized activity (%)
120
100
80
60
40
20
0
WT
CC 48-49AA
C127A
C183A
C211S
C264A
C314A
C0

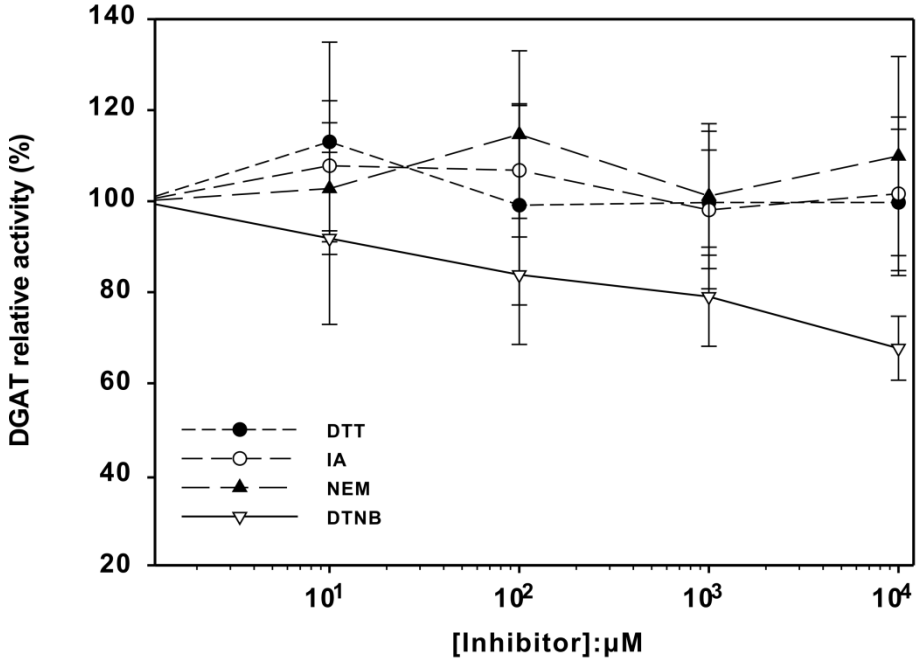


Figure 4.3. Mutant C0 is not susceptible to thiol-modifying reagents. Microsome fractions containing V5-tagged C0 mutant were treated with IA, DTNB, DTT and NEM and DGAT assay was performed under the same condition as in Figure 4.1. Enzyme activity from microsomes not treated with thiol-specific reagents was set as 100%. The calculations were preformed as described in Figure 4.1. Values represent means ± S.D for biological triplicate analysis.

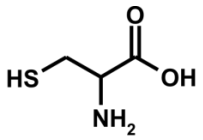
4.3.3 Cysteines are not involved in disulfide linkages

Disulfide linkage is a common structural element that facilitates proper folding of proteins to maintain structural integrity (Woycechowsky and Raines, 2000). To investigate if cysteine residues in DGAT2 participate in protein structure and folding, we attempted to map possible disulfide linkages using thiol-specific modification, which is a variant of the substituted-cysteine accessibility method (Lu and Deutsch, 2001). The application of this methodology has been reported for acyl-CoA:cholesterol acyltransferase 1 (ACAT1, EC 2.3.1.26), another enzyme involved in neutral lipid biosynthesis (Guo et al., 2005a). This strategy takes advantage of the mass-tagging competition between a large and a small thiol-alkylating reagent, which is manifested in different band shifts of the protein on an immunoblot. Under SDS-denaturing conditions, the free cysteines of ScDGAT2 were modified by combinations of different concentrations of PEG-mal (MW 5000) and IA (MW 185) (Figure 4.4A). When ScDAGT2 was only modified by IA, the mobility rate of the modified protein on SDS-PAGE did not exhibit a significant change (Lane 2, Figure 4.4B). With increasing ratio of PEG-mal to IA, a ladder of modified ScDGAT2 species appeared (Lanes 3-8, Figure 4.4B) until only PEG-mal was incubated with the protein (Lane 9, Figure 4.4B). The protein band reached the maximum shift when all the free cysteines were linked to PEG-mal molecules. A total of seven extra bands in addition to the band

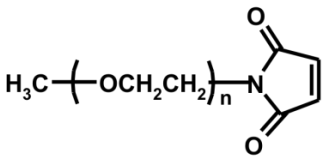
representing unmodified ScDGAT2 were identified, which indicated the presence of seven free cysteines. To verify this, we performed an indirect modification as a complementary test. In this case, IA was incubated with ScDGAT2 first to block the free cysteine residues. DTT or 2-mecaptoethanol was then used to reduce the possible bonded cysteines followed by modification with PEG-mal. As shown in Figure 4.4B (Lanes 11-12), a band shift did not occur as a result of treatments with DTT or 2-ME compared to a control implying that no cysteines were modified by PEG-mal. As an additional control, the same experiment was also carried out using the mutant C0. As expected, no band shift was appeared on immunoblots (Figure 4.4C), which confirmed that the additional bands visualized in Figure 4.4B corresponded to the modification of cysteines in ScDGAT2. These results signified that the cysteine residues in ScDGAT2 are not involved in disulfide linkage.

Figure 4.4. Analysis of disulfide bonds in ScDGAT2. (A) Chemical structure of cysteine residue and mPEG-mal. (B) Direct and indirect thiol-modification of ScDGAT2 under denatured condition. (C) Direct and indirect thiol-modification of mutant C0 as negative control. The details of modification are described under "Materials and methods." After modification, samples were used for further analysis by SDS-PAGE and immunoblotting. In lanes 1-9, the value reported at the top of each lane represented the concentration of IA and PEG-mal used in direct modification. Lanes 11-12 display the results of indirect modification and lane 10 is the negative control for indirect modification. The results are from one of two representative experiments.

A

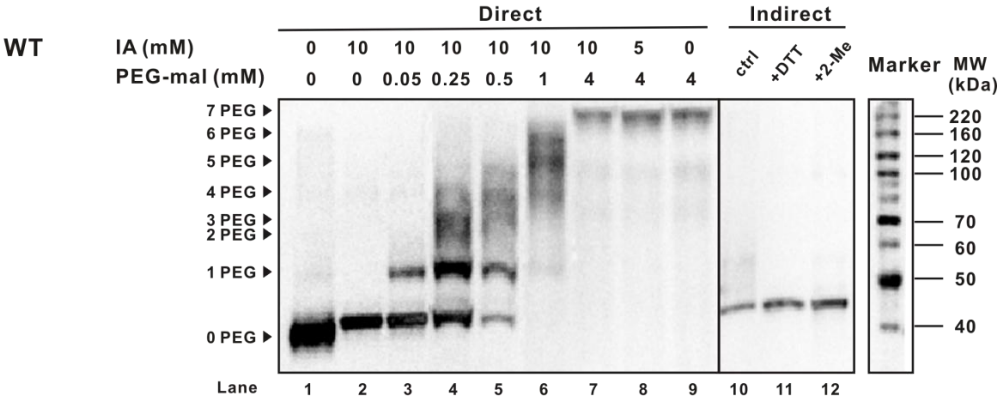


Cysteine

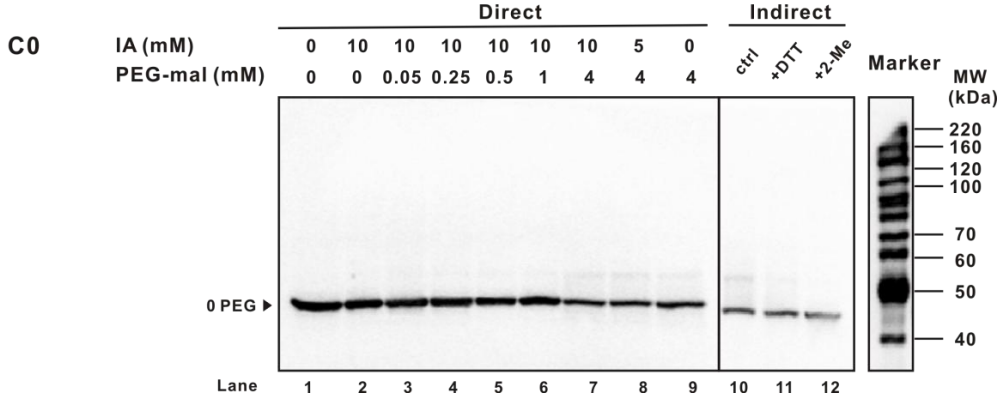


mPEG-mal MW: 5K

B



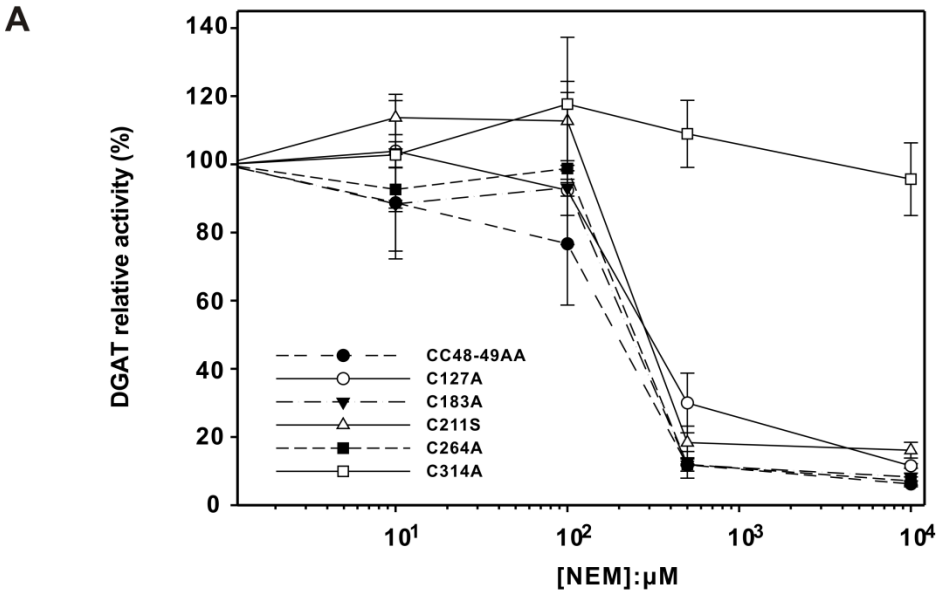
C

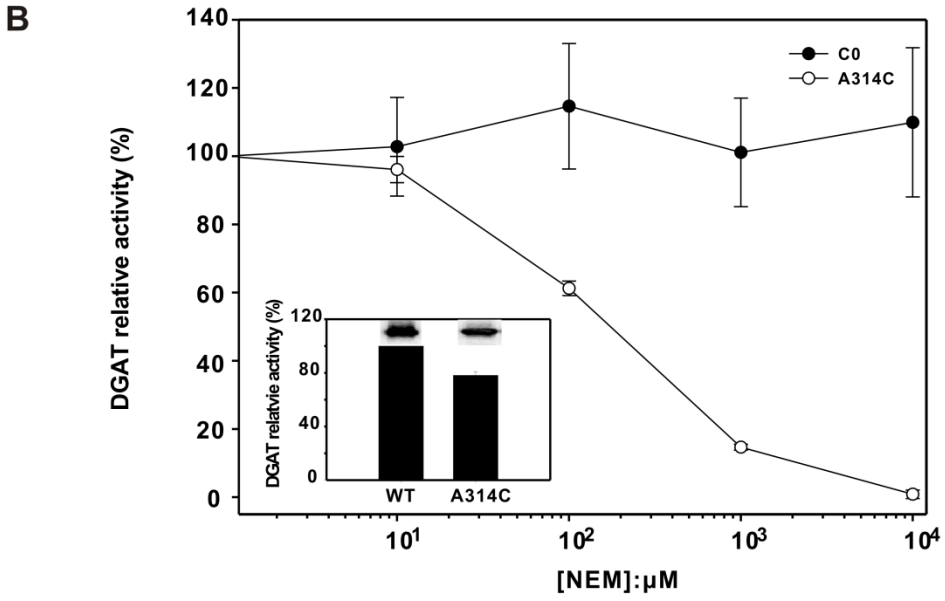


4.3.4 Cys^{314} is the locus of NEM-mediated inhibition

Since the cysteines of ScDGAT2 were not directly involved in the enzyme catalysis, we sought to further investigate the origin of the observed thiol-specific inhibition by characterizing the cysteine residue(s) responsible for mediating the inhibition of ScDGAT2 by NEM. Because there was no disulfide bond in ScDGAT2, we performed the inhibition assay with all cysteine mutants (CC48-49AA, C127A, C183A, C211S, C264A and C314A) individually. Among all mutants, only C314A retained significant DGAT activity even at the highest NEM concentration utilized (Figure 4.5A). These data demonstrate that Cys^{314} is the major target of NEM-induced inhibition of ScDGAT2 activity. To confirm these results we generated the mutant A314C using C0 as the template where a single cysteine residue was present at position 314 in ScDGAT2. Mutant A314C possessed substantial enzyme activity and expression level compared to wild type ScDGAT2 (inset of Figure 4.5B). Mutant A314C was inactivated by NEM in a similar fashion to the wild type enzyme with an IC_{50} value of 160 μM. In contrast, C0 was not sensitive to NEM, supporting the data obtained with C314A mutant (Figure 4.5B).

Figure 4.5. Sensitivity of cysteine mutants on NEM-mediated inhibition. (A) DGAT activity inhibition of ScDGAT2 cysteine mutants by NEM. Enzyme activities from mutants without NEM treatment were assigned as 100%. Values represent means ± S.D for triplicate analysis in one experiment. The results represent one of two representative experiments. (B) Inhibition of DGAT activity in mutant A314C by NEM. The values were compared with data from mutant C0. Enzyme activity and expression level of mutant A314C in contrast to wild type protein are shown in the inset. IC_{50} value was determined as described in Figure 4.1.


A
DGAT relative activity (%)
140
120
100
80
60
40
20
0
CC48-49AA
C127A
C183A
C211S
C264A
C314A
10^1
10^2
10^3
10^4
[NEM]:μM


B
DGAT relative activity (%)
140
120
100
80
60
40
20
0
C0
A314C
DGAT relatvie activity (%)
120
80
40
0
WT
A314C
10^1
10^2
10^3
10^4
[NEM]:μM

4.4 Discussion

In the present study, we report a systematic survey of the local functional and structural roles of seven cysteine residues in ScDGAT2 using chemical modification in conjunction with SDM. For membrane proteins such as DGAT2, it is a valuable approach to obtain structural and functional information by investigating the contribution of cysteine residues to protein conformation and/or catalytic mechanism such as described for ACAT1 (Lu et al., 2002).

We have demonstrated that microsomal ScDGAT activity was susceptible to inhibition by thiol modification using various reagents. Genetic modifications of ScDGAT2 through SDM, however, showed that none of the cysteine residues is essential for the enzyme catalysis. Thus, the catalytic mechanism of DGAT2 does not require the formation of a covalent acyl-thioester intermediate, as reported for other acyltransferases such as chalcone synthase (Ferrer et al., 1999). In addition, analysis of the accessibility of the sulfhydryl groups revealed that cysteines are also not involved in formation of disulfide linkages.

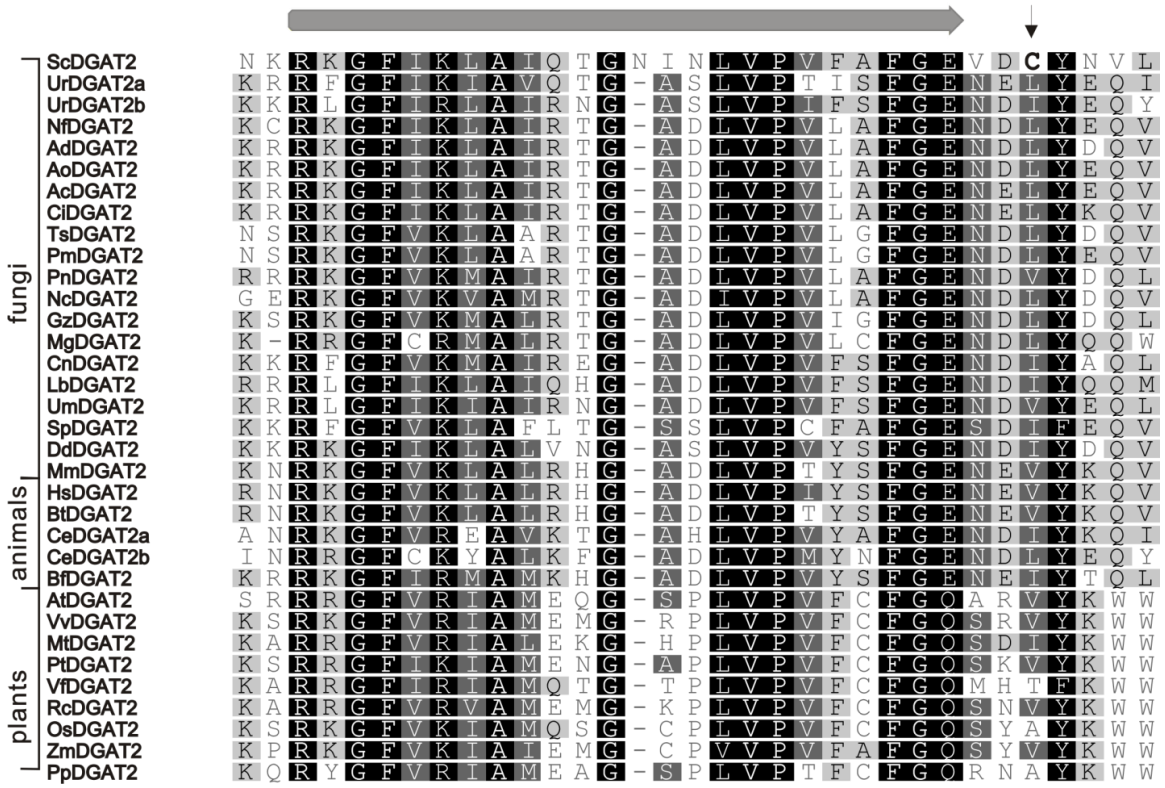
Although the cysteine residues in ScDGAT2 were not directly involved in enzyme catalysis or disulfide bond formation, the covalent attachment of a bulky group (ie.NEM) to the free sulfhydryl group of these residues likely triggered steric hindrance leading to protein inactivation. Therefore, it is possible to explain inhibition in terms of the size of attached group and concentration of thiol-specific

reagent. In addition, modification on Cys^{314} was at the center of the inhibition. Cys^{314} probably resides in a critical position which is near the active site of the enzyme. It is interesting to note that cysteine residues in ACAT1, an enzyme structurally and functionally related to DGAT1, were also non-essential (Lu et al., 2002). Cys^{467} in the C-terminal region of ACAT1 was the major target for inhibition, but in this case, the cysteine residue resided in a hydrophobic region and was close to the putative active site His^{460} (Guo et al., 2005a; Lu et al., 2002). This polar histidine was proposed to be in the active site of a superfamily of membrane-bound O-acyltransferases including enzymes such as DGAT1, ACAT and wax ester synthase (MBOAT, NCBI domain ID pfam03062), which are involved in transferring fatty acids onto membrane-embedded targets (Hofmann, 2000). The cluster of DGAT2 family proteins in the phylogenetic tree, however, is clearly different from that of MBOAT family, suggesting that DGAT2 might not share similar catalytic sites with these enzymes (Hishikawa et al., 2008). A closer look at the DGAT2 protein alignment showed that a consensus sequence RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q) (residues 288 to 311 of ScDGAT2) precedes Cys^{314} in the C-terminal region (Figure 4.6). This motif is the most conserved region in thirty-four DGAT2 sequences from a wide range of organisms including animals, plants and fungi. Residues such as arginine, lysine and glutamic acid in this region, could serve as ideal molecules to facilitate

protonation and deprotonation, which are important steps in many catalytic reactions as exemplified by the proposed catalytic mechanism of histone acyltransferase (Marmorstein, 2001). It is possible that covalent attachment of thiol-specific groups to Cys^{314}, which might be near the active site (s) in this molecular signature, triggers steric interference within the protein conformation. In future experiments, it will be valuable to validate the functional significance of these multiple residues using site-saturation mutagenesis, possibly in conjunction with a high-throughput screening system which has been applied in our studies with DGAT1 (Siloto et al., 2009b; Siloto et al., 2009c). Although our evidence for Cys^{314} being near a putative active site is compelling, the possibility still exists that Cys^{314} may be involved in protein folding in a way which is not related to disulfide bonding. Modification on Cys^{314} may have hampered the proper folding state within local regions or domains which involve interactions among some key residues in the proximity of each other (Dobson, 2003). Structural transitions occurring during protein folding can be investigated in detail by many biophysical techniques ranging from optical methods to NMR spectroscopy (Dobson et al., 1998). These methods, however, are still in their infancy for membrane proteins. Moreover, it is necessary to find effective strategies to overcome the difficulties in DGAT purification before utilizing these methodologies.

Figure 4.6. Sequence alignment of DGAT2 family members. Thirty-four DGAT2 peptides from animals, plants and fungi were aligned using the Geneious (Biomatters Ltd.) with the ClustalW algorithm, showing the relationship between most conserved domain (indicated by the bar on the top) and Cys^{314} (represented by the arrow). Gray shades denote the polarity of blocks of conserved residues. Abbreviations for each DGAT2 polypeptide indicating the organism of origin and corresponding accession numbers are as follows: Ac, *Ajellomyces capsulatus*, XP_001540241; Acl, *Aspergillus clavatus*, XP_001273210; Ao, *Aspergillus oryzae*, XP_001822244; At, *Arabisopsis thaliana*, NP_566952; Bf, *Branchiostoma floridae*, XP_002208225; Bt, *Bos Taurus*, CAD58968 Ce, *Caenorhabditis elegans*, CeDGAT2a, CAB04533, CeDGAT2b, AAB04969; Ci, *Coccidioides immitis*, XP_001240299; Cn, *Cryptococcus neoformans*, EAL20089; Dd, *Dictyostelium discoideum*, XP_635762; Gz, *Gibberella zeae*, XP_381525; Hs, *Homo sapiens*, AAK84176; Lb, *Laccaria bicolor*, EDR14458; Mg, *Magnaporthe grisea*, XP_368741; Mm, *Mus musculus*, AAK84175; Mt, *Medicago truncatula*, ACJ84867; Nc, *Neurospora crassa*, CAE76475; Nf, *Neosartorya fischeri*, XP_001261291; Os, *Oryza sativa*, NP_001057530; Pm, *Penicillium marneffei*, XP_002146410; Pn, *Phaeosphaeria nodorum*, EAT89076; Pp, *Physcomitrella patens*, XP_001777726; Pt, *Populus trichocarpa*, XP_002317635; Rc, *Ricinus communis*, AAY16324; Sc, *Saccharomyces*

cerevisiae, NP_014888; Sp, *Schizosaccharomyces pombe*, XP_001713160; Ts, *Talaromyces stipitatus*, EED21737; Um, *Ustilago maydis*, XP_760084; Ur, *Umbelopsis ramanniana*, UrDGAT2a, AAK84179, UrDGAT2b, AAK84180; Vf, *Vernicia fordii*, ABC94474; Vv, *Vitis vinifera*, CAO68497; Zm, *Zea mays*, ACG38122.


fungi
animals
plants
ScDGAT2
UrDGAT2a
UrDGAT2b
NfDGAT2
AdDGAT2
AoDGAT2
AcDGAT2
CiDGAT2
TsDGAT2
PmDGAT2
PnDGAT2
NcDGAT2
GzDGAT2
MgDGAT2
CnDGAT2
LbDGAT2
UmDGAT2
SpDGAT2
DdDGAT2
MmDGAT2
HsDGAT2
BtDGAT2
CeDGAT2a
CeDGAT2b
BfDGAT2
AtDGAT2
VvDGAT2
MtDGAT2
PtDGAT2
VfDGAT2
RcDGAT2
OsDGAT2
ZmDGAT2
PpDGAT2

In summary, we have provided insights into the local structural and functional role of cysteine residues in ScDGAT2. Although these cysteines were not directly involved in enzyme-mediated catalysis and structural support through disulfide bonding, we have demonstrated that ScDGAT2 could be inactivated by thiol-specific modification on a single cysteine in the C-terminus which might be near the active site or related to protein folding. Our insights into the role of cysteine residues in ScDGAT2 have set up the foundation for further work towards elucidating the details of structure/function in this enzyme. In addition, our approach may prove useful for investigating the role of cysteine residues in other membrane-bound enzymes.

A version of this chapter has been submitted for publication. Liu, Q., Siloto R. M. P., Snyder C. L. and Weselake R. J. 2010. Journal of Biological Chemistry (under revision).

5. Functional and Topological Characterization of Signature Motifs in Acyl-CoA:Diacylglycerol Acyltransferase 2

5.1 Introduction

In eukaryotes, triacylglycerol (TAG or TG) serves as both a source of energy and a reservoir of fatty acids for membrane formation and maintenance. Acyl-CoA:diacylglycerol acyltransferase (DGAT or DAGAT, EC 2.3.1.20) is found mainly in the endoplasmic reticulum (ER), and catalyzes the final and committed step in acyl-CoA-dependent TAG biosynthesis (Lung and Weselake, 2006; Yen et al., 2008).

There are at least two known microsomal DGAT family members DGAT1 and DGAT2, which do not share substantial sequence homology and are proposed to have distinct physiological roles in TAG metabolism. DGAT1 is a member of a superfamily of membrane-bound O-acyltransferases (MBOAT) (Hofmann, 2000), whereas DGAT2 belongs to a family including acyl-CoA:monoacylglycerol acyltransferase (MGAT) and acyl-CoA:wax alcohol acyltransferase (AWAT) (Turkish et al., 2005). Orthologues of both DGAT1 and DGAT2 have been widely identified in animals and plants. Reverse genetic studies in DGAT knock-out mice have demonstrated that DGAT2 plays a dominant role in TAG

biosynthesis in mammals (Stone et al., 2004). Down-regulation of DGAT2 expression in mice resulted in increased resistance to obesity as well as significant improvement in hepatic steatosis and insulin sensitivity (Choi et al., 2007; Liu et al., 2008; Yu et al., 2005). DGAT2 expression levels are also considered to be related to the skin disease psoriasis (Wakimoto et al., 2003). Therefore, DGAT2 represents an important therapeutic target for management or treatment of these disorders. In certain plant species, DGAT2 plays a role in the selective accumulation of unusual fatty acids into TAG (Kroon et al., 2006; Shockey et al., 2006). Overexpression of these DGAT2s in crop plants represents a potential means of producing value-added oils to meet the growing demand from industry (Snyder et al., 2009). In fungi such as the yeast *Saccharomyces cerevisiae*, DGAT2 (also known as Dga1p) is the only characterized DGAT, and is responsible for most of the TAG synthesis (Sandager et al., 2002). Yet, a detailed molecular understanding of DGAT2 function is lacking due to the intrinsic technical difficulties associated with purification and crystallization of this membrane protein (Little et al., 1994).

To date, there have been few reports providing insights into the molecular aspects of DGAT2-related enzymes. Experimental data from murine DGAT2 have led to a topology model in which DGAT2 is proposed to possess one or two adjacent transmembrane domains (TMDs) at the N-terminus, with both termini

residing in the cytosol. In this model, the bulk of the protein at the C-terminus faces toward the cytosolic side and likely contains the active site of the protein (Stone et al., 2006). Studies using tung tree DGAT2 expressed in tobacco cells also demonstrated that both the N- and C- termini were exposed to the cytosol (Shockey et al., 2006). Although the location of the putative TMDs was consistent with the predicted result, it remains unclear whether the N-terminus of the protein consists of two TMDs or one long domain embedded in the membrane bilayer. Site-directed mutagenesis (SDM) showed that two conserved regions are catalytically essential for murine DGAT2 (Stone et al., 2006). One region in the first TMD was proposed to be a putative neutral lipid-binding domain and the other region is the tetrapeptide HPHG, facing the cytosol.

As an initial step toward functional and structural analysis of this enzyme, we previously used yeast DGAT2 (ScDGAT2) as a model to probe the role of cysteines in enzyme catalysis and structural support; we found that cysteines were not essential to either of these processes, but a widely conserved motif at the C-terminus of yeast DGAT2 was identified which may include an active site or be involved in protein folding (Liu et al., 2010).

In a continuing investigation to better understand the molecular mechanism of this enzyme, we have probed the functional and structural importance of signature motifs in ScDGAT2. Through *in silico* analyses, we

uncovered unique sequence features and conserved motifs in ScDGAT2. Mutagenesis revealed that some of the studied motifs are functionally important for ScDGAT2 activity, including a previously undisclosed hydrophilic stretch found only in this enzyme. We then evaluated their topological orientation using chemical modification of single cysteines. Our work provides new experimental evidence that ScDGAT2 has a different topology compared to murine DGAT2.

5.2 Materials and methods

5.2.1 Materials

Zymolase 100T was obtained from Seikagaku Corp (East Falmouth, MA). Complete protease inhibitor cocktail tablets were purchased from Roche Molecular Biochemicals). [1-^{14}C]Oleoyl-CoA was from GE Healthcare (Baie d'Urfe, Quebec, Canada). Diolein was from Avanti Polar Lipids (Alabaster, AL). mPEG5000-maleimide (PEG-mal, 5kDa) was obtained from SunBio (Orinda, CA). Anti-V5-HRP antibodies were purchased from Invitrogen (Burlington, ON). Rabbit anti-Kar2p antibodies were from Santa Cruz Biotechnology (Redwood City, CA). The goat anti-rabbit-HRP secondary antibodies were from Zymed (San Francisco, CA).

5.2.2 Construction of ScDGAT2 mutants

The coding region for the yeast *S. cerevisiae* DGAT2 was PCR-amplified and ligated into the yeast expression vector pYES2.1-TOPO (Invitrogen) and

fused with a C-terminal V5 tag (referred to as ScDGAT-V5). N- and C- terminal truncation mutants N1(Δ1-62), N2 (Δ1-33), C1 (Δ374-418), C2 (Δ391-418), C3 (Δ413-418), C4 (Δ413-418, 413::A6) as well as N3 in which V5 tag was fused to N terminus instead, were similarly constructed. Other deletion, insertion and point mutants including TM (Δ70-91), DS1 (Δ150-187), DS2 (Δ156-174), DS3 (162::163-172), DS4 (162::GA5), H193A, H195A, EPHS (H193E/G195S), YFP (Y129A/F130A/P131A), F71A and L73A were constructed by PCR-driven overlapping extension according to the method of Heckman and Pease (2007) using ScDGAT2-V5 as a template. Plasmids coding for various single Cys-containing mutants were generated by the same method using a previously constructed Cys-less mutant (C0) of ScDGAT2-V5 as the template, in which all seven endogenous cysteines were mutated to alanines (Liu et al., 2010). A list of primers used for the construction of various mutants is given in Table 5.1. A schematic representation of each mutant is also shown in supplemental Figure 5.1. All constructs were sequenced to verify their integrity and that the genes were cloned in-frame with the epitope tag.

Table 5.1. Oligonucleotides used for mutations within ScDGAT2

Mutation	Forward (Fwd) and reverse (Rev) oligonucleotides in 5' to 3' orientation
N1 (Δ1-62)	Fwd: TCAGGGATCCGACTATGGCTGTAGCATGGCACACTTCTTCATTTG Rev: CCCAACTATCTTCAATTCTGCATCCGGTAC
N2 (Δ1-33)	Fwd: GCGATGATCGATAAAAGGGAACAGACTCTCAAACC Rev: CCCAACTATCTTCAATTCTGCATCCGGTAC
N3	Fwd:ATATGGATCCGCGATGGGTAAGCCTATCCCTAACCCTTTACTCGGTCTCG ATTCTACGTCAGGAACATTCAATGATATAAGAAGAAGGAAGAAGGAAGAAG GAAGCCCTACAGCC Rev: ATATAAGCTTTTACCCAACTATCTTCAATTCTGCATCCGGTACCCCATATTTTT CTC
C1 (Δ374-418)	Fwd: TCAGGGATCCGACCATGTCAGGCACTTTCAATGATATAAG Rev: TTATAAGCTTGACGTATATAGGCCTTCCAACAACAAC
C2 (Δ391-418)	Fwd: TCAGGGATCCGACCATGTCAGGCACTTTCAATGATATAAG Rev: TTTAAGCTTGTCATGGAAATGATTAACAACATC
C3 (Δ413-418)	Fwd: TCAGGGATCCGACCATGTCAGGCACTTTCAATGATATAAG Rev: TTTAAGCTTTTTTTCTCTATTTTCGTAAT
C4 (Δ413-418, 413::A_6)	Fwd: TCAGGGATCCGACCATGTCAGGCACTTTCAATGATATAAG Rev:TATAAGCTTCGCCGCCGCTGCTGCTGCTGCATCCGGTACCCCATATTTTTC TCTATTTTCGTAATATAGTCT
TM1 (Δ70-91)	Fwd: GATAGGTCTCCTGCAACTGGCGAAGTGG Rev: TGCAGGAGACCTATCAGAAGTGTGCCATGCTACAGCCAGAGTTTG
DS1 (Δ150-187)	Fwd: ACAGGACCAACGTACTTATTCGGTTACCATCC Rev: CGTTGGTCCTGTCGTGAAAGTTGGCTTTAAATTGAC
DS2 (Δ156-174)	Fwd: ACTATTGACTATCGGAACCAGGAATGTACAGGGCC Rev: GGTTCCGATAGTCAATAGTTTTTTCGTTAACTCTC
DS3(162::163-172)	Fwd: CCTACAAAATACAGTATAAACTTAAAGTCTCCAACTAAGTATTCCATTAATCT CAAAAGC Rev: AGACTTTAAGTTTATACTGTATTTTGTAGGCCACAATCTAATCTTGTAATTTTT TTC
DS4 (162::GA_5)	Fwd:GGTGCTGGTGCTGGTGCTGGTGCTGGTGCTCCAACTAAGTATTCCATTAA TCTCAAAAGCAACTCTACTATTGACTATCGCAA Rev:AGCACCAGCACCAGCACCAGCACCAGCACCCCACAATCTAATCTTATAA TTTTTTTCGTTAACTCTCTTATTTTTTGAAAGCGTAAA
H193A	Fwd: GTACTTATTTGGTTACGCGCCGCACGGCATAGGAGCAC Rev: GTGCTCCTATGCCGTGCGGCGCGTAACCAAATAAGTAC
H195A	Fwd: CTTATTTGGTTACCATCCGGCCGGCATAGGAGCACTTGG Rev: CCAAGTGCTCCTATGCCGGCCGGATGGTAACCAAATAAG
EPHS	Fwd: TACTTATTCGGATACGAGCCACACTCCATAGGAGCAC Rev: GTATCCGAATAAGTACGTTGGCCCTGTACATTC
YFP	Fwd: CATTTGGAAGTGGTATTGTGATGCTGCCGCTATAAGTTTGATTAAAAC Rev: GTTTTAATCAAACTTATAGCGGCAGCATCACAATACCACTTCCAAATG
F71A	Fwd: GGCACACTTCTTCAGCTGTACTCTTCTCC

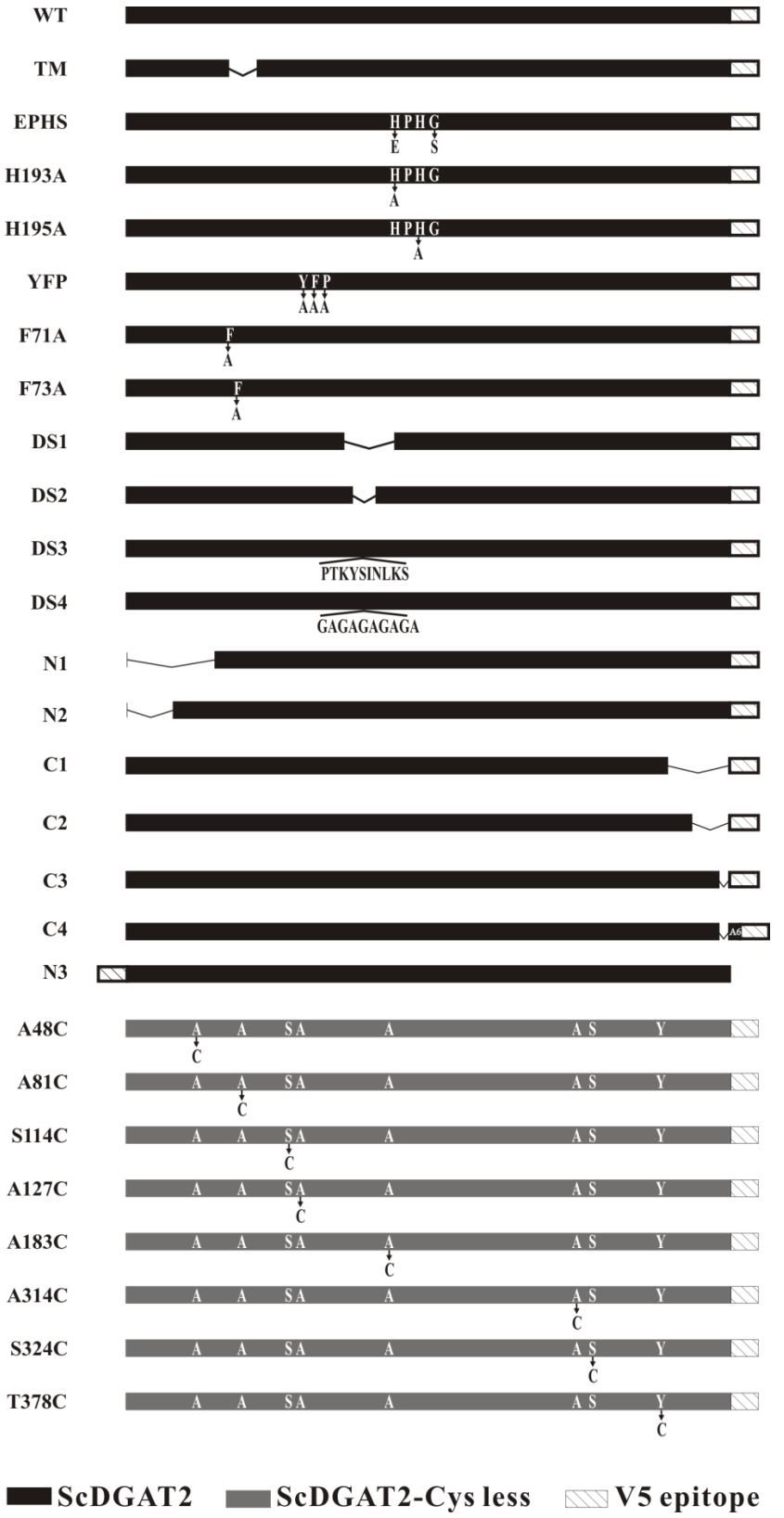
Table 5.1. Continued

	Rev: GGAGAAGAGTACAGCTGAAGAAGTGTGCC
L73A	Fwd: GCACACTTCTTCATTTGTAGCCTTCTCCATATTTACG Rev: CGTAAATATGGAGAAGGCTACAAATGAAGAAGTGTGC
A48C*	Fwd: ACCACAACTAGAGTCATGCGCTCCATTGGC Rev: GCCAATGGAGCGCATGACTCTAGTTGTGGT
A81C*	Fwd: CATATTTACGTTATTTTGCATCTCGACACCAG Rev: CTGGTGTCGAGATGCAAAATAACGTAAATATG
S114C*	Fwd: AACCGATACTGTCTTCGATTTCG Rev: CGAAATCGAAGACAGTATCGGTT
A127C*	Fwd: GGAAGGCCTATATGCGTTGAAAAGAAAATA Rev: TATTTTCTTTTCAACGCATATAGGCCTTCC
A183C*	Fwd: TCGCAACCAGGAATGTACAGGGCCAACG Rev: CGTTGGCCCTGTACATTCCTGGTTGCGA
A253C*	Fwd: TCGGAAAAACTGTTTAAGGACTC Rev: GAGTCCTTAAACAGTTTTTCCGA
A314C*	Fwd: TTGGAGAGGTGGACTGTTATAATGTTCTGAGC Rev: GCTCAGAACATTATAACAGTCCACCTCTCCAA
S324C*	Fwd: AAAAAAGATTGTGTCCTGGGTAA Rev: TTACCCAGGACACAATCTTTTTT
T378C*	Fwd: GTTGAAAAGAAAATATGCAATCCGCCAGATGAT Rev: ATCATCTGGCGGATTGCATATTTTCTTTTCAAC

* Template used was Cys-less ScDGAT2 mutant.

Underline below the sequence denotes the codon where site-directed mutagenesis was applied.

Figure 5.1. Schematic design of ScDGAT2 mutants. The figure shows the position of the epitope tag as well as point, insertion and deletion mutations within wild type (WT) ScDGAT2 (rectangular and black block) and Cys-less ScDGAT2 (rectangular and gray block).


WT
TM
EPHS
H P H G
E S
H193A
H P H G
A
H195A
H P H G
A
YFP
Y F P
A A A
F71A
F
A
F73A
F
A
DS1
DS2
DS3
PTKYSINLKS
DS4
GAGAGAGAGA
N1
N2
C1
C2
C3
C4
A6
N3
A48C
A A SA A A S Y
C
A81C
A A SA A A S Y
C
S114C
A A SA A A S Y
C
A127C
A A SA A A S Y
C
A183C
A A SA A A S Y
C
A314C
A A SA A A S Y
C
S324C
A A SA A A S Y
C
T378C
A A SA A A S Y
C
ScDGAT2
ScDGAT2-Cys less
V5 epitope

5.2.3 Strains and culture conditions

A yeast quadruple knockout strain H1246 (*MATα are1*-Δ*::HIS3, are2*-Δ*::LEU2, dga1*-Δ*::KanMX4, lro1*-Δ*::TRP1 ADE2*), which is devoid of DGAT activity and TAG biosynthesis, was used as the host strain for protein expression experiments (provided by Dr. S. Stymne and U. Ståhl) (Sandager et al., 2002). Constructs were transformed into the yeast cells using the lithium acetate/single-stranded carrier DNA/PEG method (Gietz and Schiestl, 2007). Yeast cells were also transformed with an empty vector or pYESLacZ as a control. The transformants were selected on plates lacking uracil and cultivated in yeast nitrogen base with 2% (w/v) dextrose (YNBD) medium. Cells were then harvested, washed with water and inoculated in induction medium where dextrose was replaced by 2% (w/v) galactose and 1% (w/v) raffinose to induce the gene expression.

5.2.4 Yeast microsome preparation

Right side-out microsomes were prepared for topological analysis as described elsewhere with following modifications (Gilstring and Ljungdahl, 2000). Cells were collected, washed with water and resuspended in spheroplasting buffer (1.2M soribitol, 50mM potassium acetate, 20mM Hepes, pH 7.5, 2mM dithiothreitol). Cells were then converted to spheroplasts by incubating at 30 ^{o}C for 15min in the presence of 0.75mg/mL zymolase 100T. The spheroplasts were

collected by centrifugation at 1000×*g* for 5min at 4°C and resuspended in lysis buffer (100mM sorbitol, 50mM potassium acetate, 20mM Hepes, pH 7.4) supplemented with complete protease inhibitor cocktail. Spheroplasts were lysed by 20 strokes in a Dounce homogenizer. Unlysed cells and cell debris were pelleted by a second centrifugation at 1000×*g*. After another centrifugation at 12,000x*g* for 10min, the supernatant was subjected to centrifugation at 100,000×*g* for 30min at 4°C to recover the microsomes. The resulting pellets were resuspended in 10mM Hepes-NaOH, pH 7.4.

5.2.5 DGAT assay

Equivalent amounts of microsomal protein were prepared and used for *in vitro* DGAT activity assays as previously described in Chapter 3. Briefly, DGAT activity was measured as the production of [1-^{14}C]TAG from [1-^{14}C]oleoyl-CoA and unlabeled dioleoyl-DAG in a 10min assay. TAG was separated from the reaction mixture by TLC and the radioactivity was quantified using a Beckman-Coulter LS6500 liquid scintillation counter.

5.2.6 Fractionation and extraction of membrane proteins

Preparation of right side-out microsomes was performed as described in 5.2.4. An equal volume of 2% SDS (w/v) or 0.2 M Na_2CO_3 (pH 11) was added to the suspended microsomal proteins. After incubation at 4°C for 1h, samples were centrifuged at 100,000×*g* for 30min at 4°C to separate the supernatant and pellet

fractions. The pellets were brought up to the same volume as the supernatant with resuspension buffer (10mM Hepes-NaOH, pH 7.4). Samples from supernatant and pellet were mixed with SDS loading buffer and subjected to SDS-PAGE and immunoblotting.

5.2.7 Thiol-specific chemical blocking and modification

Thiol-specific chemical blocking with N-ethylmaleimide (NEM) followed by modification with PEG-mal under non-denaturing and denaturing conditions was performed using modifications of the procedures described by Wang (Wang et al., 2008). For modification under non-denaturing conditions, microsomal protein was diluted to 1mg/mL and aliquoted equally for one of three treatments: 1) 1h incubation on ice with 0.5mM PEG-mal (to modify cysteines facing the cytoplasm), followed by addition of dithiothreitol (DTT) to 10mM and further incubation for 10min to quench the reaction; 2) 30min incubation with 5mM NEM to block accessible suflhydryl groups, followed by quenching with DTT as above; 3) untreated control.

For modification under denaturing conditions, diluted microsomal vesicles were divided into two aliquots. One was directly treated with 0.5mM PEG-mal in the presence of 2% SDS followed by addition of DTT and incubation for 10min as described above. The other half was first incubated with 5mM NEM on ice for 30min in the absence of SDS followed by immediate ultracentrifugation

(100,000×*g*, 30min). The pellet was washed once with 1mL of ice-cold resuspension buffer to remove excess NEM. After another ultracentrifugation, the pellet was resuspended in 50μL resuspension buffer. The sample was then incubated with 0.5mM PEG-mal in the presence of 2% SDS at room temperature for 20min. The reaction was finally quenched by incubation with 10mM DTT for 10min. All aliquots were mixed with sample buffer and subjected to SDS-PAGE.

5.2.8 SDS-PAGE and immunoblotting analysis

Protein samples were resolved by 10% SDS-PAGE and transferred to a PVDF membrane. V5-tagged ScDGAT2 mutants were detected using anti-V5-HRP antibodies according to the protocol described by the manufacturer. Control protein chaperone Kar2p was detected using anti-Kar2p as the primary antibodies followed by incubation with goat anti-rabbit-HRP secondary antibodies. The bound antibodies were detected using the ECL detection system (GE Healthcare). Densitometry analysis of immunoblots was performed using ImageQuant TL software (GE Healthcare).

5.2.9 Computational methods

The following algorithms were used for predicting the topological organization of DGAT2: HMMTOP (Tusnady and Simon, 2001), TMHMM (Krogh et al., 2001), TopPred (Claros and Vonheijne, 1994), SVMtm(Yuan et al., 2004), SOSUI (Hirokawa et al., 1998), Conpred II (Arai et al., 2004), TMpred

(Hofmann and Stoffel, 1993). Sequence alignments were conducted with Geneious (Biomatters Ltd.) using the ClustalW algorithm. Sequence alignments were performed using homologous DGAT2 polypeptides obtained by BLAST search. Sequences from the different organisms with corresponding accession numbers are as follows: Ac, *Ajellomyces capsulatus*, XP_001540241; Acl, *Aspergillus clavatus*, XP_001273210; Ao, *Aspergillus oryzae*, XP_001822244; At, *Arabidopsis thaliana*, NP_566952; Bf, *Branchiostoma floridae*, XP_002208225; Bt, *Bos taurus,* CAD58968 Ce, *Caenorhabditis elegans*, CeDGAT2a, CAB04533, CeDGAT2b, AAB04969; Ci, *Coccidioides immitis,* XP_001240299; Cn, *Cryptococcus neoformans*, EAL20089; Dd, *Dictyostelium discoideum*, XP_635762; Gz, *Gibberella zeae*, XP_381525; Hs, *Homo sapiens*, AAK84176; Lb, *Laccaria bicolor*, EDR14458; Mg, *Magnaporthe grisea*, XP_368741; Mm, *Mus musculus*, AAK84175; Mt, *Medicago truncatula*, ACJ84867; Nc, *Neurospora crassa*, CAE76475; Nf, *Neosartorya fischeri*, XP_001261291; Os, *Oryza sativa*, NP_001057530; Pm, *Penicillium marneffei*, XP_002146410; Pn, *Phaeosphaeria nodorum*, EAT89076; Pp, *Physcomitrella patens*, XP_001777726; Pt, *Populus trichocarpa*, XP_002317635; Rc, *Ricinus communis*, AAY16324; Sc, *Saccharomyces cerevisiae*, NP_014888; Sp, *Schizosaccharomyces pombe*, XP_001713160; Ts, *Talaromyces stipitatus*, EED21737; Um, *Ustilago maydis*, XP_760084; Ur, *Umbelopsis ramanniana*,

UrDGAT2a, AAK84179, UrDGAT2b, AAK84180; Vf, *Vernicia fordii*, ABC94474; Vv, *Vitis vinifera*, CAO68497; Zm, *Zea mays*, ACG38122.

5.3 Results

5.3.1 Sequence motifs in ScDGAT2

Multiple sequence alignment is an essential tool for predicting important structural features of proteins. To identify DGAT2 conserved motifs *in silico*, thirty-four representative polypeptides from a wide range of organisms were aligned using Clustal W algorithm. Three regions with remarkable sequence conservation were located (Figure 5.2). The motif YFP (^{129}YFP^{131} in ScDGAT2) is completely conserved in DGAT2 from different organisms with exception of castor bean, where it corresponds to HFP. Second, the motif HPHG (^{193}HPHG^{196} in ScDGAT2) is conserved in sequences from animals and fungi and is essential for DGAT2 activity in mice (Stone et al., 2006). In plant DGAT2s, the corresponding motif is EPHS. A third motif consisting of RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q) spans the longest conserved region in DGAT2-homologous proteins. This motif is found between residues 288 and 311 in ScDGAT2 and is in close proximity to Cys^{314}, which has been previously demonstrated to be the locus of NEM-mediated enzyme inhibition (Liu et al., 2010). It is noteworthy that a putative neutral lipid-binding domain, FLXLXXX^{n} (where n is a non-polar amino acid), earlier reported in murine

DGAT2 (Stone et al., 2006) could not be identified in ScDGAT2. Instead, the corresponding motif, FVLF, is found between positions 71 and 74. The sequence alignment also showed a segment of 38 amino acids situated between YFP and HPHG motifs (positions 150 to 187 in ScDGAT2) which is unique to ScDGAT2 and absent in plant, animal and other fungal DGAT2s.

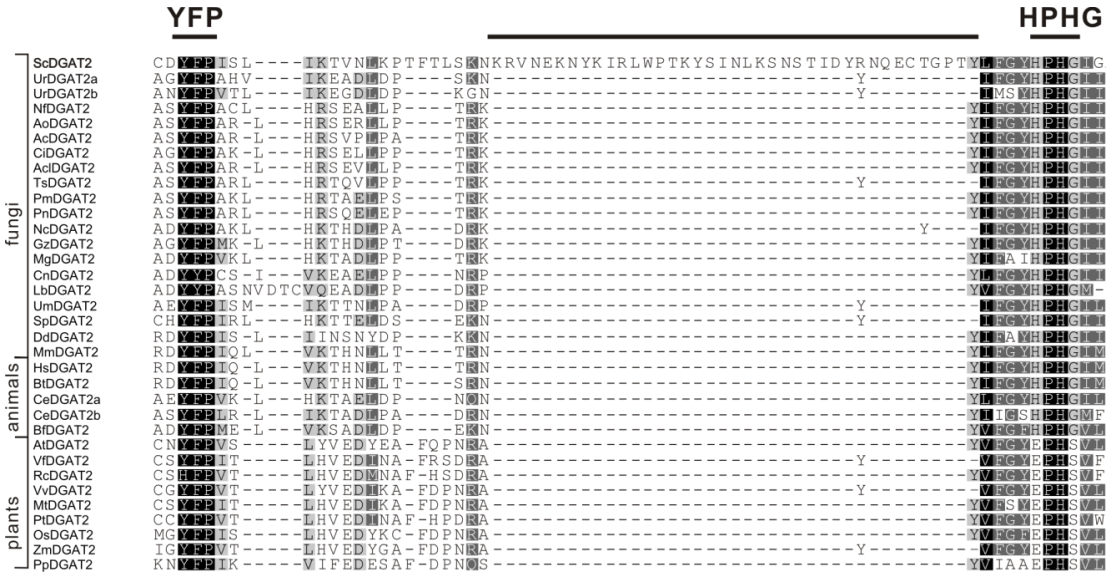


Figure 5.2. Sequence motifs in DGAT2 family. Thirty-four DGAT2 polypeptides from animals, plants and fungi were aligned by Geneious (Biomatters Ltd.). The conserved motif YFP and HPHG are indicated. Abbreviations for each DGAT2 polypeptide indicating the organism of origin and corresponding accession numbers are described under "Material and Methods."

5.3.2 Hydropathy profile analysis and TMD prediction of DGAT2s

To further explore the potential structural aspects using in silico approaches, we evaluated the hydropathy profile and topology distribution of ScDGAT2 in comparison to murine DGAT2 whose topology was experimentally tested (Stone et al., 2006). The distribution of sequence motifs in yeast and murine DGAT2 was compared in Kyte-Doolittle (Kyte and Doolittle, 1982) hydropathy plots (dotted boxes, Figure 5.3). Both proteins show similar plots, especially at N-terminal region, which contains a hydrophilic stretch followed by a hydrophobic region. ScDGAT2, however, contains segments that are more hydrophobic than murine DGAT2, especially between residues 192 and 250 which have a contiguous sequence of mostly hydrophobic amino acids (Figure 5.3A). In addition, the 38 residue segment found exclusively in ScDGAT2 (Figure 5.3A) resides in a hydrophilic region and is clearly positioned between the conserved motif ^{129}YFP^{131} and ^{193}HPHG^{196}.

We also subjected both DGAT2 sequences to seven common topological prediction programs based on different algorithms (Table 5.2). Murine DGAT2 was predicted to have one or two abutting TMDs located at the N-terminal region, and an N-terminus facing the cytosol, which is in agreement with the experimentally determined topology. In the case of ScDGAT2, the orientation of N-terminus was generally predicted to be towards the cytosol with three to four TMDs in the polypeptide. The first TMD (residues 66 to 97) was predicted by all

of the algorithms and is comparable with the model of murine DGAT2. There was more variability, however, in the predictions for the location of putative TMDs in the region between residues 188 to 236. Two additional TMDs were predicted between amino acids 290-313 and 341-359. These predictions corroborate the hydropathy plot analysis suggesting that ScDGAT2 may have a different topology compared to murine DGAT2, which could affect the orientation of the signature motifs discussed above.

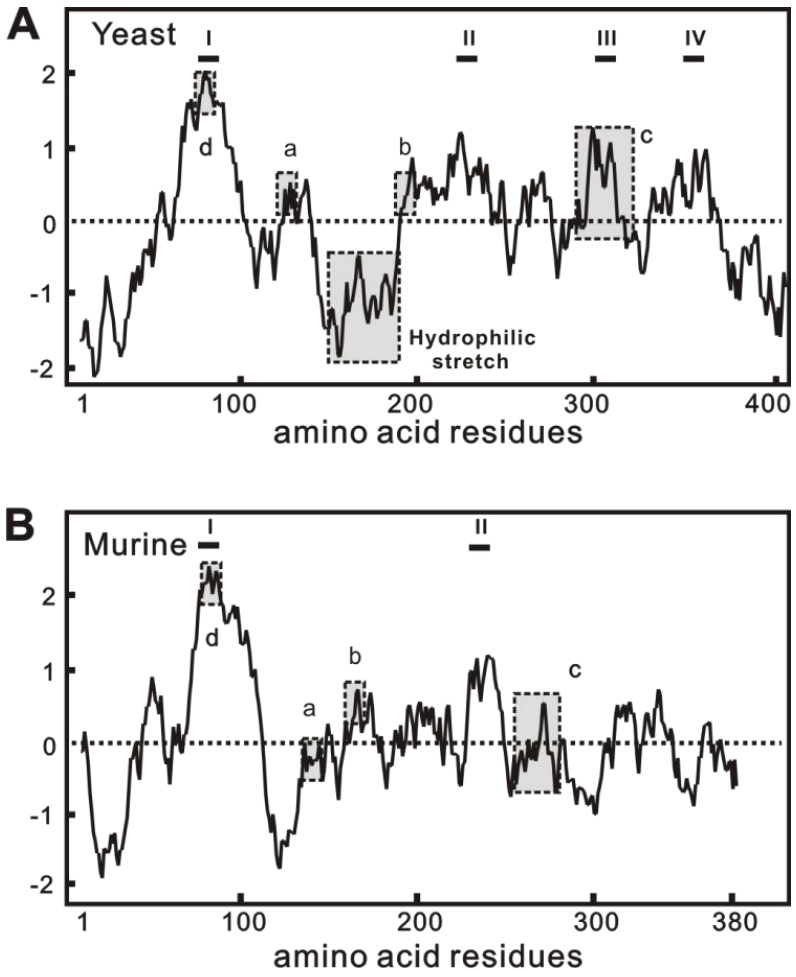


Figure 5.3. Hydropathy analysis of yeast and murine DGAT2s. Aligned hydropathy plots of yeast (A) and murine (B) DGAT2 proteins were generated by the method of Kyte and Doolittle (Kyte and Doolittle, 1982) using a window size of 19. The most hydrophobic segments which may contain putative TMDs are indicated by Roman numerals. Motifs studied here are boxed and labeled as follows: a (YFP), b (HPHG), c (RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q) and d (putative lipid-binding site).

Table 5.2. Predicted transmembrane domains (TMDs) in yeast and murine DGAT2s. The polypeptides corresponding to yeast and murine DGAT2s were submitted to seven transmembrane prediction algorithms.

	HMMTOP	TMHMM	TopPred	SVMtm	SOSUI	ConPred II	TMpred
ScDGAT2	68-92	70-92	68-88	66-97	75-97	73-93	72-92
	—	—	—	193-207	188-210	189-209	196-214
	200-224	—	215-235	217-231	213-234	215-235	216-236
	294-310	—	290-310	—	—	293-313	—
	341-359	—	—	—	—	—	—
Number of TMD	4	1	3	3	3	4	3
Orientation*	IN	IN	IN	N/A	N/A	IN	IN
MmDGAT2	68-92	73-95	76-96	73-93	66-88	70-90	76-96
	—	—	—	95-109	93-115	—	—
	—	—	230-250	—	—	—	224-248
Number of TMD	1	1	2	2	2	1	2
Orientation*	IN	IN	IN	N/A	N/A	IN	IN

*Orientation of N-terminus. Cytosol-"IN".

5.3.3 Site-directed mutagenesis of signature motifs

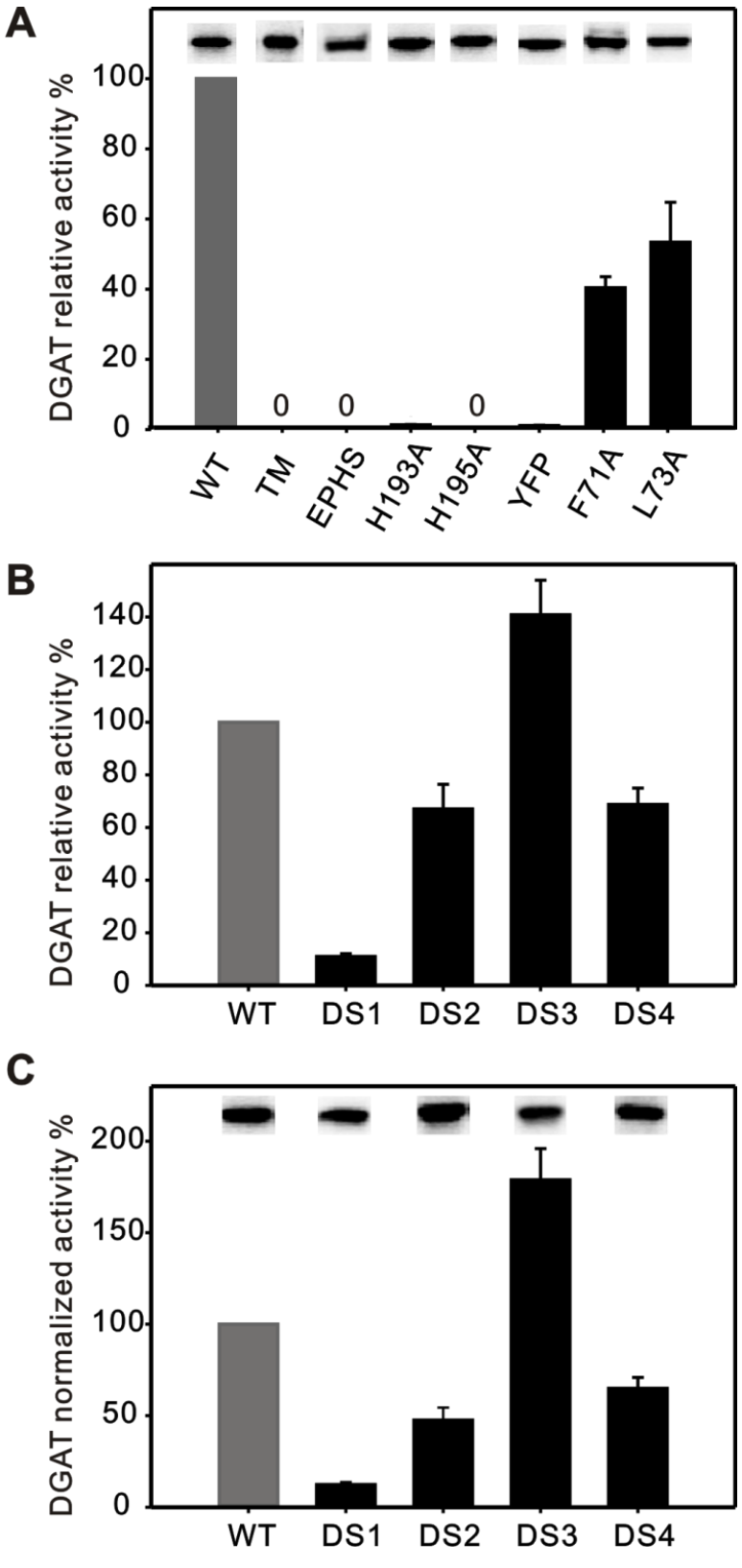
To determine the functional significance of these motifs in ScDGAT2, we conducted SDM using recombinant in *S. cerevisiae* strain H1246. The impact of these mutations was evaluated by measuring the in vitro enzyme activities and protein expression using isolated microsomes. Because the strain H1246 is devoid of DGAT activity, results are a direct measurement of the performance of each mutant. The expression of wild type (WT) ScDGAT2 and ensuing mutants was verified using immunoblotting by taking advantage of the C-terminal V5 tag. As shown in Figure 5.4A, substitution of ^{129}YFP^{131} to ^{129}AAA^{131} (mutant YFP) and H193 to A193 (mutant H193A) led to almost complete loss of enzyme activity. Similarly, substitution of H195 to A195 and replacement of ^{193}HPHG^{195} to corresponding motif EPHS found in plant DGAT2s (mutants H195A and EPHS, respectively) both abolished the enzyme activities. Deletion of the first putative TMD between residues 70 and 91 also resulted in the total loss of activity (mutant TM).The F^{71} to A^{71} and L^{73} to A^{73} substitution in the sequence corresponding to the putative lipid-binding motif in murine DGAT2 (mutant F71A and F73A, respectively) retained significant activity (over 40%) compared to the WT ScDGAT2 . Immunoblotting analysis of microsomes used for the enzyme assays revealed similar accumulation of each mutant protein in comparison to WT (insets

of Figure 5.4A), indicating that the observed changes in activity were not a result of differential expression levels.

The fact that the hydrophilic stretch between motif ^{129}YFP^{131} and ^{193}HPHG^{196} (Figure 5.3A) is unique for ScDGAT2 raised the question of whether this segment plays a role in enzyme function. To address this, we performed deletion and insertion mutagenesis on this segment. Deletion of the whole segment (mutant DS1, Δ150-187), resulted in an enzyme with less than 20% of WT activity (Figure 5.4B) whereas removal of a smaller segment (mutant DS2, Δ156-174) preserved over 60% of the original activity. In view of these results, we hypothesized that insertions in this region might have a positive effect on enzyme activity. Thus, we duplicated a stretch of 10 amino acids from resides 163 to 172 at position 162 creating mutant DS3 (162::163-172). Interestingly, this mutant showed a 40% increase in enzyme activity compared to the WT. To investigate whether this effect was specific to the sequence duplicated, we inserted a decapeptide composed of $(GA)_5$ in the same position. In contrast, the resulting mutant, (DS4, 162::GAGAGAGAGA) presented a decrease of activity of about 30% compared to WT suggesting that the positive effect on the activity in DS3 was sequence-related. This trend was retained after normalizing activity to protein expression levels (Figure 5.4C), confirming that the observed changes in activity did not result from variation of protein expression levels. As expected, no

DGAT activity was detected in H1246 yeast cells transformed with empty vector or the control pYESLacZ (data not shown). Collectively, these results show that the unique hydrophilic stretch in ScDGAT2 may play a key role in modulating enzyme activity.

Figure 5.4. Site-directed mutagenesis of signature motifs. (A) *In vitro* DGAT activities of mutants TM (Δ70-91), H193A, H195A, EPHS (H193E/G195S), YFP (Y129A/F130A/P131A), F71A and L73A. Microsomes from yeast cells expressing mutants and WT protein were incubated in the presence of diolein and [1-^{14}C]oleoyl-CoA and assayed as described under "Material and Methods." Activities are expressed as picomoles of TAG formed per minute per milligram of microsomal protein and quoted relative to the activity of WT protein (set to 100%). Data represent the average ± standard deviation (n=3). Immunoblotting analysis demonstrating the expression level of mutants is shown in the insets. (B) Relative activities of mutants DS1 (Δ150-187), DS2 (Δ156-174), DS3 (162::163-172) and DS4 (162:: GA_5). Values represent average ± the standard deviation (n=3). (C) Normalized activities of mutants DS1, D2S, DS3 and DS4. The DGAT relative activity was normalized by expression level of the protein which was measured through comparative densitometry analyses of immunoblots. Data represent the average with the standard deviation (n=3). Immunoblots of the microsomal protein prepared from cells expressing these mutants are shown in the insets.


A
DGAT relative activity %
100
80
60
40
20
0
0
0
0
WT
TM
EPHS
H193A
H195A
YFP
F71A
L73A
B
DGAT relative activity %
140
120
100
80
60
40
20
0
WT
DS1
DS2
DS3
DS4
C
DGAT normalized activity %
200
150
100
50
0
WT
DS1
DS2
DS3
DS4

5.3.4 ScDGAT2 is an integral membrane protein

Our previous studies have indicated that ScDGAT2 can be copiously detected in microsomal membrane fraction (Liu et al., 2010) which was also observed in the immunoblots from Figure 5.4A. Hydropathy analysis and topology prediction suggest that ScDGAT2 is likely an integral membrane protein and probably has a different topology compared to murine DGAT2. Before exploring topological details, we sought to experimentally determine the nature of membrane association of ScDGAT2. Microsomes containing V5-tagged ScDGAT2 were extracted with either detergent (SDS) or alkaline salt (Na_2CO_3) and then separated by ultracentrifugation into supernatant and pellet which were analyzed by immunoblotting. As shown in Figure 5.5 (lane 3 and 4), ScDGAT2 could be extracted from the membrane and was present in the supernatant only when microsomes were treated with 1% SDS detergent, a condition that disrupts the integrity of membrane bilayers. In contrast, when the membrane was treated with 0.1M Na_2CO_3, an alkaline salt commonly used to strip peripheral membrane proteins, also releasing luminal proteins, DGAT2 remained in the pellet (Figure 5.5, lane 5 and 6). The same experiment was conducted using Kar2p, a well-characterized chaperone present in ER lumen (Normington et al., 1989). As expected, this control protein was detected in the soluble fractions after treatment

with either SDS or Na_2CO_3 (Figure 5.5, lane 3 to 6). These results clearly showed that ScDGAT2 is strongly associated with microsomal membranes.

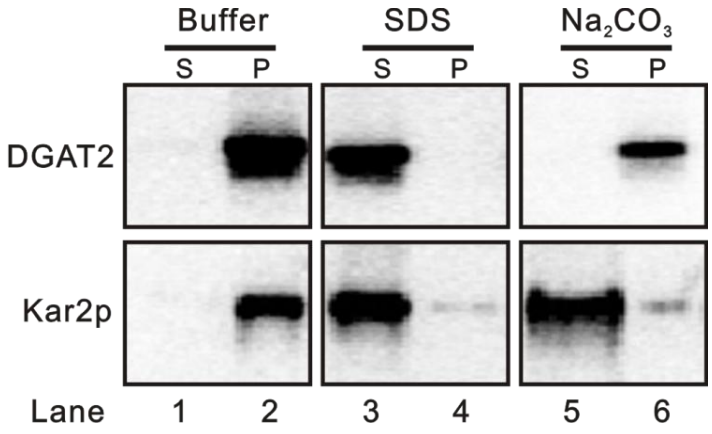


Figure 5.5. Yeast DGAT2 is an integral membrane protein. Microsomes were prepared from yeast cells expressing ScDGAT2-V5 (see "Materials and methods"). The membrane was incubated with buffer (10mM Hepes-NaOH, pH 7.4) alone (lane 1 and 2) or buffer containing 1% SDS (lane 3 and 4) or 0.1 M Na_2CO_3 at pH 11 (lane 5 and 6). After incubation, samples were separated by ultracentrifugation into the supernatant (S) or pellet (P) which was subjected to immunoblotting analysis with anti-V5 or anti-Kar2p antibodies.

5.3.5 Integrity and sidedness of microsomal vesicles

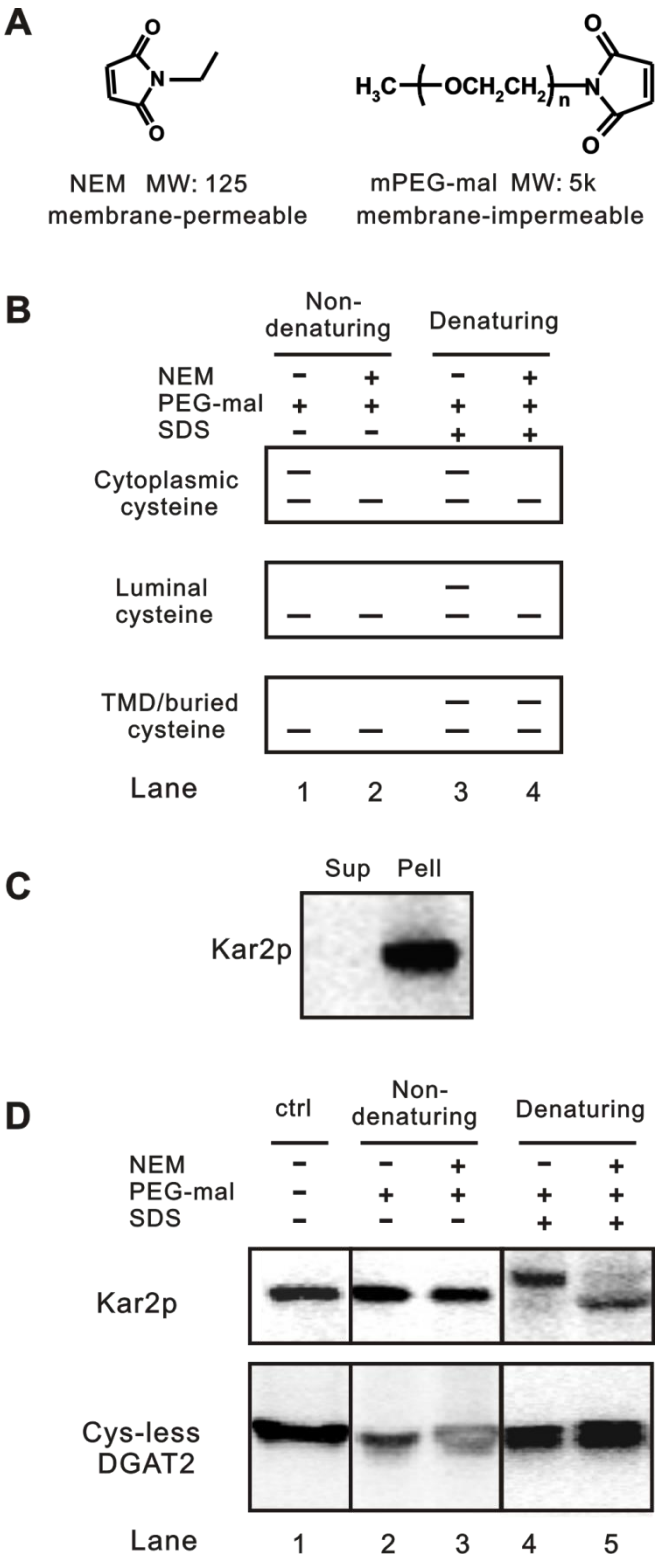
Topology is a fundamental aspect of the structure of membrane proteins. Determination of the topological orientation of signature motifs in ScDGAT2 is an important step towards understanding the catalytic mechanism of this protein. Previously, we constructed a Cys-less ScDGAT2 in which all seven native cysteine residues were substituted with alanines. This mutant retained significant activity and was expressed at comparable levels to the WT (Liu et al., 2010), indicating that it likely has a similar structure to the native ScDGAT2. These findings provide the basis for mapping membrane topology by probing the accessibility of reintroduced single cysteine residues to thiol-modifying reagents. Modification of each single cysteine can be detected by distinct band-shift patterns on an immunoblot by using thiol-specific reagents with different membrane permeabilities and molecular weights (ie., NEM and PEG-mal in Figure 5.6A). The position of reintroduced single cysteine residues relative to a TMD can yield three different immunoblot patterns under four modification conditions, as schematically represented in Figure 5.6B. If the target cysteine is exposed to the cystosol, PEG-mal, which has a large and membrane-impermeable carbohydrate moiety, can be attached, leading to a detectable band-shift (approximately 5kDa) in both SDS denaturing and non-denaturing conditions (Figure 5.6B, lanes 1 and 3). This band-shift will not appear if NEM, which has

smaller size (125 Da), is used to block the cysteine prior to PEG-mal modification (Figure 5.6B, lanes 2 and 4). Detection of a cysteine exposed to the luminal side is characterized by a band-shift with PEG-mal treatment which occurs only after membranes are disrupted under SDS denaturing condition (Figure 5.6B, lanes 1 and 3). This shift, however, is prevented by prior exposure of protein to NEM, which can permeate membranes and thus block cysteine residues on the luminal side (Figure 5.6B, lanes 2 and 4). Because the reaction between the NEM and the ionized form of thiol requires a water molecule as a proton acceptor, NEM is virtually unreactive to cysteine residues embedded in the membrane bilayer or at a protein-protein interface (Bogdanov et al., 2005; Tamura et al., 2001). This property can be used to detect buried cysteine residues that can be characterized by a band-shift when NEM is removed following PEG-mal treatment in the presence of SDS (Figure 5.6B, lane 4).

This methodology, originally developed from the gel-shift assay described by Lu and Deutsch (2001), has been successfully applied for membrane proteins in yeast (Wang et al., 2008). One major assumption with use of this approach is that topological orientation of microsome vesicles is identical to that in ER *in vivo*. To test this hypothesis, we examined the accessibility of the ER luminal protein Kar2p, which contains a single native cysteine at the N-terminus, under four different conditions depicted in Figure 5.6B. Immunoblotting analysis of cell

fractions showed that Kar2p was detected only in the microsomal fraction (Figure 5.6C), indicating that the vesicles remained sealed during preparation and can be used for accessibility assay. When the yeast microsomes were incubated with PEG-mal in the presence or absence of NEM under non-denaturing conditions, a band-shift was not observed (Figure 5.6D, lane 2 and 3), indicating that Kar2p was not exposed to the cytosol. Under denaturing conditions with SDS, a band-shift resulting from the PEG-mal attachment could be detected (Figure 5.6D, lane 4). The band representing unmodified protein was not evident, suggesting that efficient modification by PEG-mal occurred. A prior treatment with NEM substantially blocked PEG-mal modification under denaturing conditions (Figure 5.6D, lane 5), agreeing with the expected immunoblot pattern for the luminal cysteine residue in Figure 5.6B. The same treatment was repeated with Cys-less ScDGAT2 as an additional control (Figure 5.6D) and, as expected, no band-shifts were detected. Thus, we validated the thiol-modification methodology, showing also that the microsomal vesicles retained correct orientation.

Figure 5.6. Thiol-specific modification to map membrane topology. (A) Chemical structures and properties of NEM and PEG-mal. (B) Schematic depiction of immunoblot patterns from the cysteine accessibility assay. (C) Immunoblot of yeast Kar2p. Microsome pellet (pell) was separated from cytosolic protein (sup) in supernatant by 100,000×*g* centrifugation. Equal amounts of protein from each fraction were resolved by 10% SDS-PAGE and analyzed by immunoblotting with an anti-Kar2p antibody. (D) Membrane integrity and orientation as determined by thiol-specific modification. Microsomal membranes were prepared as detailed under "Materials and methods" and treated or not with NEM under non-denaturing conditions, followed by treatment with PEG-mal in the presence or absence of SDS. In the condition that NEM-treated sample was used for PEG-mal modification in the presence of SDS, the NEM was removed. The reaction mixtures were analyzed by immunoblotting using anti-Kar2p and anti-V5 antibodies to detect Kar2p and Cys-less DGAT2, respectively. The results are representative of two independent experiments.

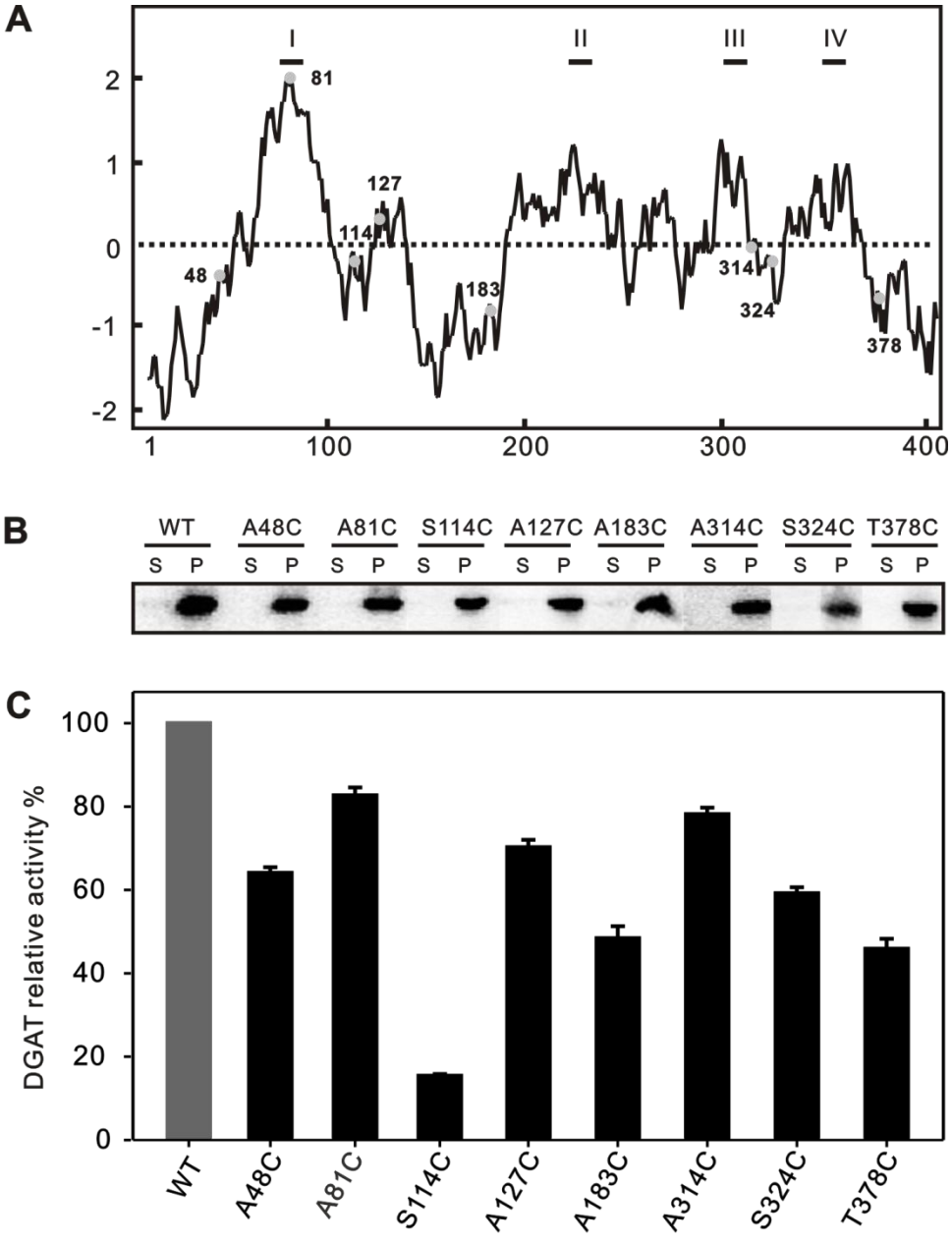

A
NEM MW: 125
membrane-permeable
mPEG-mal MW: 5k
membrane-impermeable
B
Non-denaturing
Denaturing
NEM
PEG-mal
SDS
Cytoplasmic cysteine
Luminal cysteine
TMD/buried cysteine
Lane
1 2 3 4
C
Sup Pell
Kar2p
D
ctrl
Non-denaturing
Denaturing
NEM
PEG-mal
SDS
Kar2p
Cys-less DGAT2
Lane
1 2 3 4 5

5.3.6 Activities of single Cys-containing mutants

Having demonstrated the utility of the accessibility assay for mapping membrane topology, we constructed a series of single-cysteine mutants of ScDGAT2. To minimize structural perturbation, we individually re-introduced three cysteine residues back to their original positions in the native ScDGAT2 using the Cys-less version as the template, creating mutants A48C, A127C and A183C. In addition, the mutants A81C, S114C, S324C and T378C were created to evaluate topological organization of motifs that are distant from a native cysteine. In this case, substitutions were performed in positions having low conservation yet residing in close proximity to the desired motif, mitigating possible structural disturbance. The N- and C- termini were located by using mutants A48C and T378C, respectively. Mutants A81C, S114C, A127C and A183C were used to map the topological organization of ^{129}YFP^{131} and ^{193}HPHG^{196} motifs as well as the unique hydrophilic stretch, respectively. Mutant A81C served as a control because it resides in the most hydrophobic segment (position 66 to 97), universally predicted to be a TMD (Table 5.2). Together with the previously constructed mutant A314C (Liu et al., 2010), S324C was created to determine the orientation of the highly conserved domain preceding Cys314. The position of each introduced cysteine residue in the hydropathy plot is shown in Figure 5.7A and Figure 5.1.

The constructs encoding the single-cysteine residue forms of ScDGAT2 were expressed in yeast strain H1246 as previously described. To determine if replacement of a single cysteine residue produced deleterious effects on the protein folding or structure, the ScDGAT expression level and DGAT activity for each mutant were measured. The immunoblotting analysis suggested all single-Cys proteins were expressed at levels comparable to WT, localized exclusively in the microsomal fraction with the expected molecular mass of approximately 48 kDa (Figure 5.7B). In addition, results from *in vitro* enzyme assays showed that all mutants were functional, although S114C retained only around 20% activity compared to the WT (Figure 5.7C). Therefore, we concluded that the effect of cysteine residue replacement in each mutant is tolerated and does not cause gross structural changes in the protein.

Figure 5.7. Expression of various single Cys-containing mutants. (A) Hydropathy plot of ScDGAT2 showing the position and hydrophobicity environments of each introduced cysteine residue (represented by filled circles with the positions numbered). (B) Immunoblot showing expression of each mutant. The microsomal pellet (P) was separated from the cystosolic protein in the supernatant (S) by 100,000×*g* centrifugation. Equal amounts of protein from each fraction were separated by 10% SDS-PAGE and probed with anti-V5 antibodies. (C). *In vitro* DGAT activities of the single Cys-containing mutants and wild-type (WT) ScDGAT2. All enzyme activities were measured as described under "Materis and methods." The activities expressed as picomoles of TAG formed per minute per milligram of microsomal protein and quoted relative to the activity of WT protein (defined as 100%). Data represent means ± S.D. (n=3).

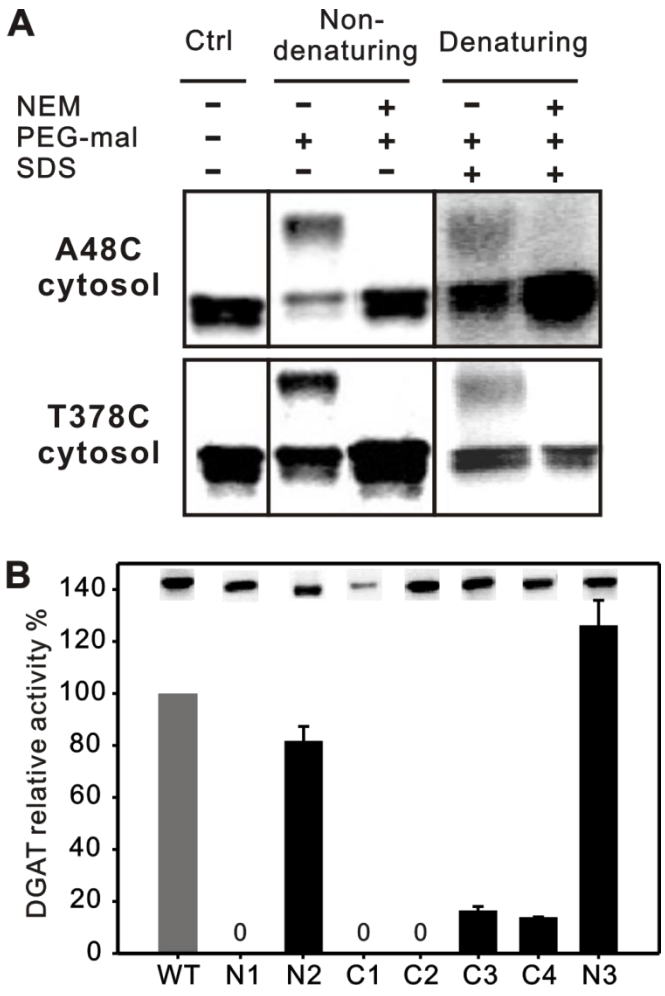

A
I
II
III
IV
81
127
114
48
183
314
324
378
2
1
0
-1
-2
1
100
200
300
400
B
WT
A48C
A81C
S114C
A127C
A183C
A314C
S324C
T378C
S P S P S P S P S P S P S P S P S P
C
DGAT relative activity %
100
80
60
40
20
0
WT
A48C
A81C
S114C
A127C
A183C
A314C
S324C
T378C

5.3.7 N- and C- termini are located in the cytosol and have different roles in enzyme catalysis

Microsomes were prepared from H1246 cells expressing A48C and T378C and subjected to thiol-specific modification and accessibility analysis as described in Figure 5.6B. The results showed that PEG-mal can modify both mutants without NEM blocking under non-denaturing or denaturing conditions (Figure 5.8A), indicating a cytoplasmic labeling pattern for both termini of ScDGAT2. To examine the role of the N- and C-terminal tails in DGAT2 function, deletion mutagenesis was carried out. Hydrophilic tails of N- and C-termini are composed of 62 and 44 amino acid residues, respectively (Table 5.2 and Figure 5.3A). Sequential deletion was performed to remove the entire N terminus (N1) and the first 33 amino acid residues (N2). Similar deletions along the C terminus were also performed to remove the entire hydrophilic tail (C1), the last 28 amino acids (C2), and the last six amino acid residues (C3). The impact of these mutations was evaluated by measuring the enzyme activity *in vitro* and protein expression level using isolated microsomes. Recombinant proteins were expressed with a C-terminal V5 tag to facilitate the detection in immunobloting. As demonstrated in Figure 5.8B, the mutant lacking the entire hydrophilic N-terminus (N1) was devoid of activity while maintaining a substantial expression level. Removal of the first 33 amino acid residues in N-terminus (N2) resulted in

minor decrease in enzyme activity (Figure 5.8B). However, deletion of the last six amino acid residues from the C-terminus (C3) caused a decrease in the enzyme activity of more than 80%. Deletion of the whole C-terminus (C1) completely abolished the enzyme activity and had a substantial impact on the protein accumulation (inset of Figure 5.8B); some expression was recovered in the C2 mutant in which about half of the hydrophilic C-terminus was deleted, but this mutant also exhibited a complete loss of enzyme activity. To further investigate the importance of the last six amino acid residues, we replaced them with six alanine residues to generate mutant C4. This mutant retained similar activity and expression level to C3, suggesting that the effect of these variations is related to the last 6 residues in the C-terminus. Considering the recombinant protein used in the current study was tagged with a C-terminal V5 epitope, we decided to investigate the influence of capping the C-terminus by instead fusing the epitope to the N-terminus (mutant N3). The enzyme activity of this mutant was approximately 20% higher compared to the WT in which the epitope was attached at the C-terminus. These results indicate that the hydrophilic C-terminal tail of ScDGAT2 is more sensitive to modifications than the N-terminus.

Figure 5.8. Sidedness and mutagenesis analysis of N- and C- termini. (A) Accessibility of mutant A48C and T378C to thiol-specific modification. Modification was performed under the same conditions described in Figure 5.6. After modification, samples were subjected to immunoblotting analysis using anti-V5 antibodies. The results shown are from one of two independent experiments. (B) DGAT activities of DGAT2 WT and mutants N1 (Δ1-62), N2 (Δ1-33), C1 (Δ374-418), C2 (Δ391-418), C3 (Δ413-418), C4 (Δ413-418, 413::A6) and N3 (N-terminus V5 tag and free C-terminus). All activities are quoted relative to that of the WT protein (set to 100%). Values represent means ± the standard deviation of experiments conducted in triplicates. Immunoblotting analysis of the microsomes prepared from cells expressing these mutants is shown in the insets.


A
Ctrl
Non-denaturing
Denaturing
NEM
PEG-mal
SDS
A48C cytosol
T378C cytosol
B
DGAT relative activity %
140
120
100
80
60
40
20
0
0
0
0
WT
N1
N2
C1
C2
C3
C4
N3

5.3.8 Topological analysis of signature motifs

The previous results indicated that both termini are located in the cytosol, which is in agreement with earlier studies on mammalian and plant DGAT2s (Shockey et al., 2006; Stone et al., 2006). *In silico* analyses, however, suggested the existence of additional TMDs in ScDGAT2. To further investigate the topological orientation of signature motifs in ScDGAT2, the accessibility of single cyseine residues in mutants A81C, S114C, A127C and A183C was analyzed. Thiol-modification experiments showed a buried modification pattern for mutant A81C which was characterized by band-shifts with both treatments under denaturing conditions (Figure 5.9). This result indicates a TMD that is likely spanning residues 66 to 97 which includes the sequence ^{71}FVLF^{74} corresponding to the putative lipid-binding site in murine DGAT2 (Stone et al., 2006). For mutant S114C, a luminal modification pattern was observed, which was evidenced by a band shift only under denaturing conditions in the absence of NEM. Although mutant A127C showed a buried modification pattern, it is unlikely that a TMD resides between positions 114 and 183, as it was not predicted with any of algorithms used in Table 5.2. In fact, only short hydrophobic stretches with low hydrophobic index are present in this region (Figure 5.3A). This is also supported by the fact that a lumenal modification pattern was observed for mutant A183C. Therefore, these data indicate that motif

^{129}YFP^{131} is in the lumen, probably in close contact with the membrane bilayer or in a highly structured region of the protein. Taken together, these results indicate that the hydrophilic segment after the first TMD is an ER luminal loop which contains motif ^{129}YFP^{131}.

Topology predictions indicate two putative TMDs from position 290 to 313 and 341 to 359 (Table 5.2). To map this region, PEG-mal accessibility assays were conducted with mutants A314C and S324C. As shown in Figure 5.9, both cysteine residues in these mutants face the cytosolic side, excluding the possibility of a TMD from residue 341 to 359, which was predicted by only one algorithm (HMMTOP). The shifted bands representing the mutant protein modified by PEG-mal were weak, indicating that residues 314 and 324 could be in close contact with the membranes or protected in a protein-protein interface with limited accessibility to PEG-mal. Taking into account *in silico* analyses (Table 5.2 and Figure 5.3A), these results suggest that the conserved motif ^{288}RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q)^{311} preceding C^{314} is the most likely region to be a TMD, returning the C-terminus of the polypeptide to the cytosol.

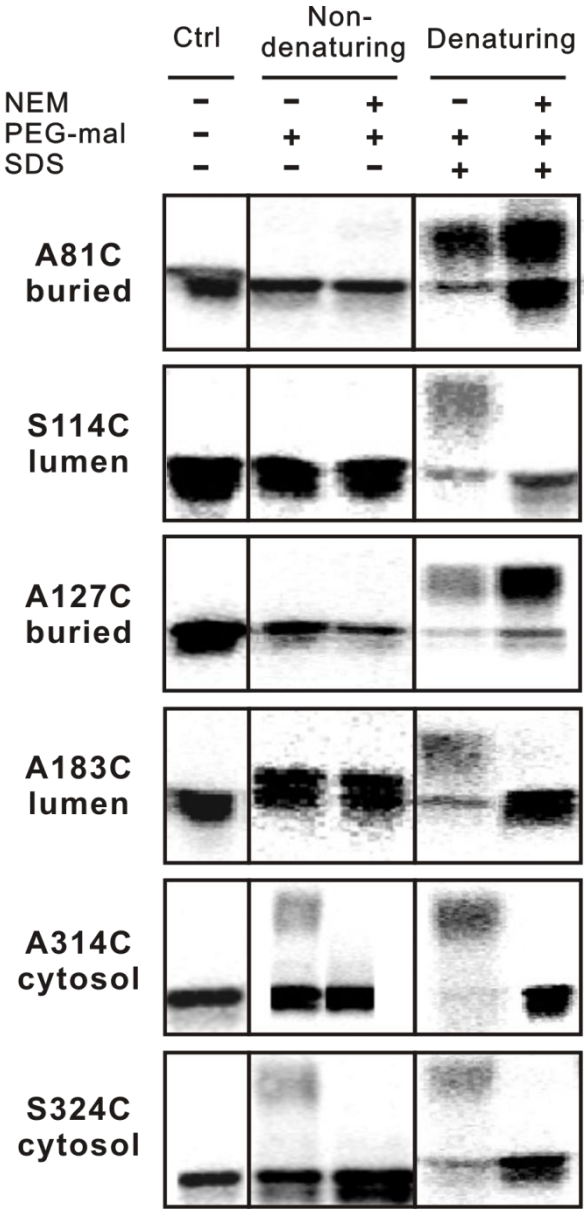


Figure 5.9. Sidedness of signature motifs. Accessibility of single Cys-containing mutant A81C, S114C, A127C, A183C, A314C and S324C to thiol-modifying reagents. Accessibility assay was conducted and detected as described previously (Figure 5.6). The results shown are from one of two independent experiments.

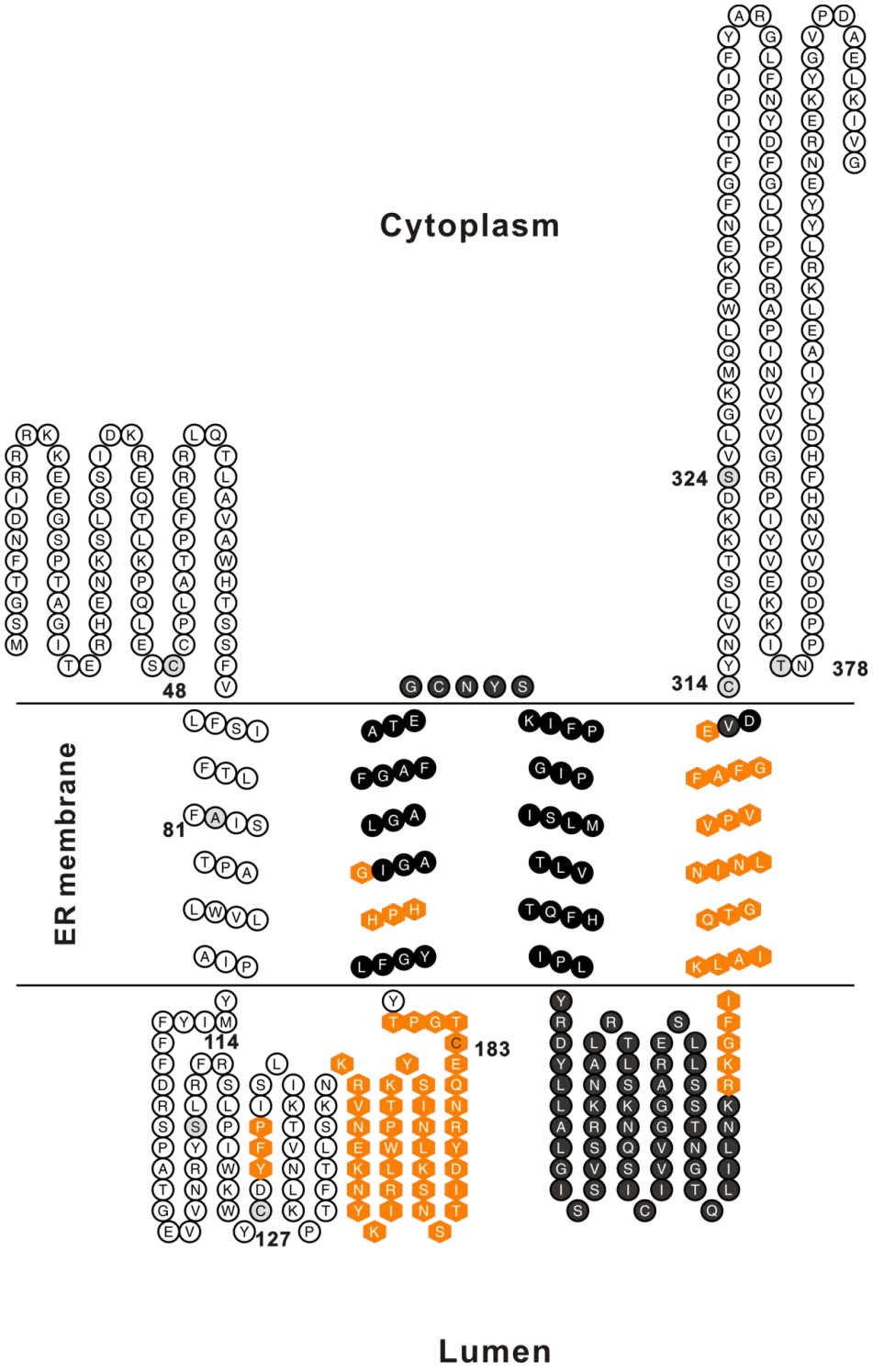
5.4 DISCUSSION

In the absence of three-dimensional structural data for DGAT2-related proteins, low-resolution structural and functional information obtained by biochemical approaches such as chemical modification in conjunction with mutagenesis can aid in the interpretation of structure-function relationships of this class of enzyme. Using these methodologies, we have determined the functional importance and topological orientation of sequence motifs in ScDGAT2.

5.4.1 Proposed model and topological distribution of important signature motifs in ScDGAT2

Based on experimental data and *in silico* analyses, our current view of the topological distribution of signature motifs in ScDGAT2 is shown in Figure 5.10. In this model, ScDGAT2 is proposed to have four TMDs with both termini oriented toward the cytosol, satisfying experimental and computational constraints. The residues highlighted in black (from the second to fourth TMD) indicate the orientation of regions established by computational analyses, whereas the other regions were supported by experimental evidence. The exact borders of TMDs were predicted using Conpred II, which applies a consensus technique and is reported to have high prediction accuracy (Persson, 2006).

Figure 5.10. Proposed topology model of ScDGAT2. This model proposes that ScDGAT2 contains four TMDs with both termini orientated toward the cytosol. Important functional motifs are represented by filled hexagons. Black-filled residues represent regions where topology was established based on *in silico* analysis only. Topology of other regions was experimentally determined. Positions where single cysteines were introduced are numbered.


Cytoplasm
ER membrane
Lumen
48
81
114
127
183
314
324
378

Both termini of yeast DGAT2 have the same sidedness, but the C-terminus may play a more important role in maintaining enzyme activity (Figure 5.8). We also observed similar results in plant homologue of DGAT1 (unpublished data). An ER retrieval motif LKLEI has been previously identified in the extreme C-terminus of tung tree DGAT2 (Shockey et al., 2006). The similar motif ^{413}ELKIVG^{418} is also present in the corresponding region of ScDGAT2 as shown in the multiple sequence alignment (Figure 5.11). Deletion or replacement of this sequence resulted in significant loss of enzyme activity, which could be related to this ER retrieval motif. Subcellular targeting experiments would help to determine whether ^{413}ELKIVG^{418} constitutes an ER retrieval motif for ScDGAT2.

It is worth noting that the N-terminus was not essential for enzyme catalysis in either murine DGAT2 (Stone et al., 2009) or ScDGAT2, and similar results were recently reported for murine DGAT1 (McFie et al., 2010). A mitochondrial targeting signal in N-terminal region of murine DGAT2 which is conserved in mammalian DGAT2s, however, was characterized between amino acid 61 to 66 . Although this signal sequence could not be identified in ScDGAT2, we cannot exclude the possibility of role of N-terminus in subcellular localization in yeast protein.

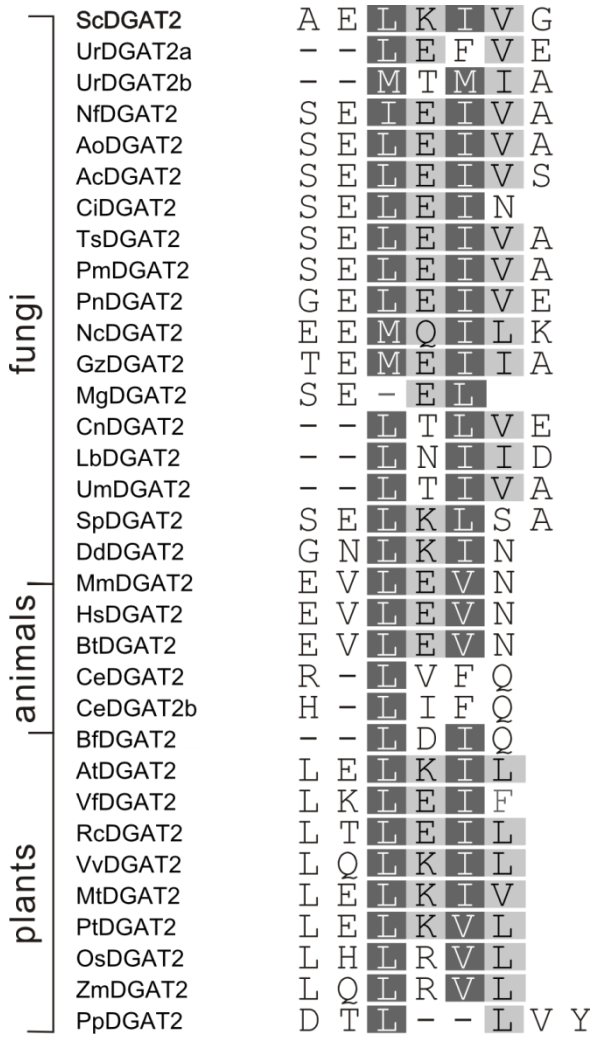


Figure 5.11. Multiple sequence alignment showing the putative ER retrieval motif in DGAT2s. Thirty-four DGAT2 polypeptides were aligned as described in Figure 5.2. Accession numbers of these protein sequences and their abbreviations suggesting the organism of origin are detailed under "Materials and methods."

The first TMD likely spans from residue 66 to 97, followed by an ER luminal loop which contains ^{129}YFP^{131} and the distinctive hydrophilic stretch (from position 150 to 187) that is essential for enzyme catalysis (Figure 5.4). Following this luminal region, three TMDs were predicted between residues 188 and 313. The fact that position 314 and 324 are likely close to a TMD strengthens the possibility of the existence of a TMD from residues 290 to 313. Therefore, the conserved motif, RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q), preceding Cys^{314} could possibly be embedded in membrane bilayer as indicated in the model. The second TMD begins with residue 188 as predicted by most of the algorithms, and contains the strongly conserved domain ^{193}HPHG^{196}. Substitution of H^{195} abolished the enzyme activity, indicating that it could be directly involved with the active site of the enzyme. This finding is consistent with the previous result from murine DGAT2 in which mutation of this histidine residue also led to a substantial loss in enzyme activity (Stone et al., 2006). A similar result was found with murine DGAT1 where substitution of H^{426} abolished enzyme activity (McFie et al., 2010). In addition, an invariant histidine residue in a DGAT1-related enzyme, acyl-CoA:cholesterol acyltransferase 1 (ACAT, EC 2.3.1.26), has also been proposed to function as a base to facilitate enzyme catalysis in the hydrophobic membrane bilayer (Guo et al., 2005b). Because of the variations in

predicting TMDs in this region, we proposed it could comprise two adjacent TMDs or one long segment embedded in the membrane bilayer.

5.4.2 Differences in topological and functional features between the yeast and murine DGAT2

A model consisting of a single hydrophobic segment embedded in the membrane bilayer with both termini exposed to cytosol has been proposed for murine DGAT2 (Stone et al., 2006). ScDGAT2 also has a characteristic hydrophobic segment near the N-terminus and two hydrophilic tails exposed to the cytosol. But our work reveals important differences in the topology of yeast and murine DGAT2, supported by several lines of computational and experimental evidence.

A relatively low level of sequence homology and differences in hydrophobic profiles of certain regions were observed between murine and yeast DGAT2s. ScDGAT2 has more hydrophobic segments and was predicted to have more TMDs compared to murine DGAT2 (Figure 5.3 and Table 5.2). In addition, deletion of the first TMD did not affect the association of ScDGAT with microsomal membranes, supporting the idea that other segments in yeast DGAT2 could possibly mediate interaction with membrane bilayers (Figure 5.4A). Furthermore, the region following the first TMD, from position 114 to 183, was

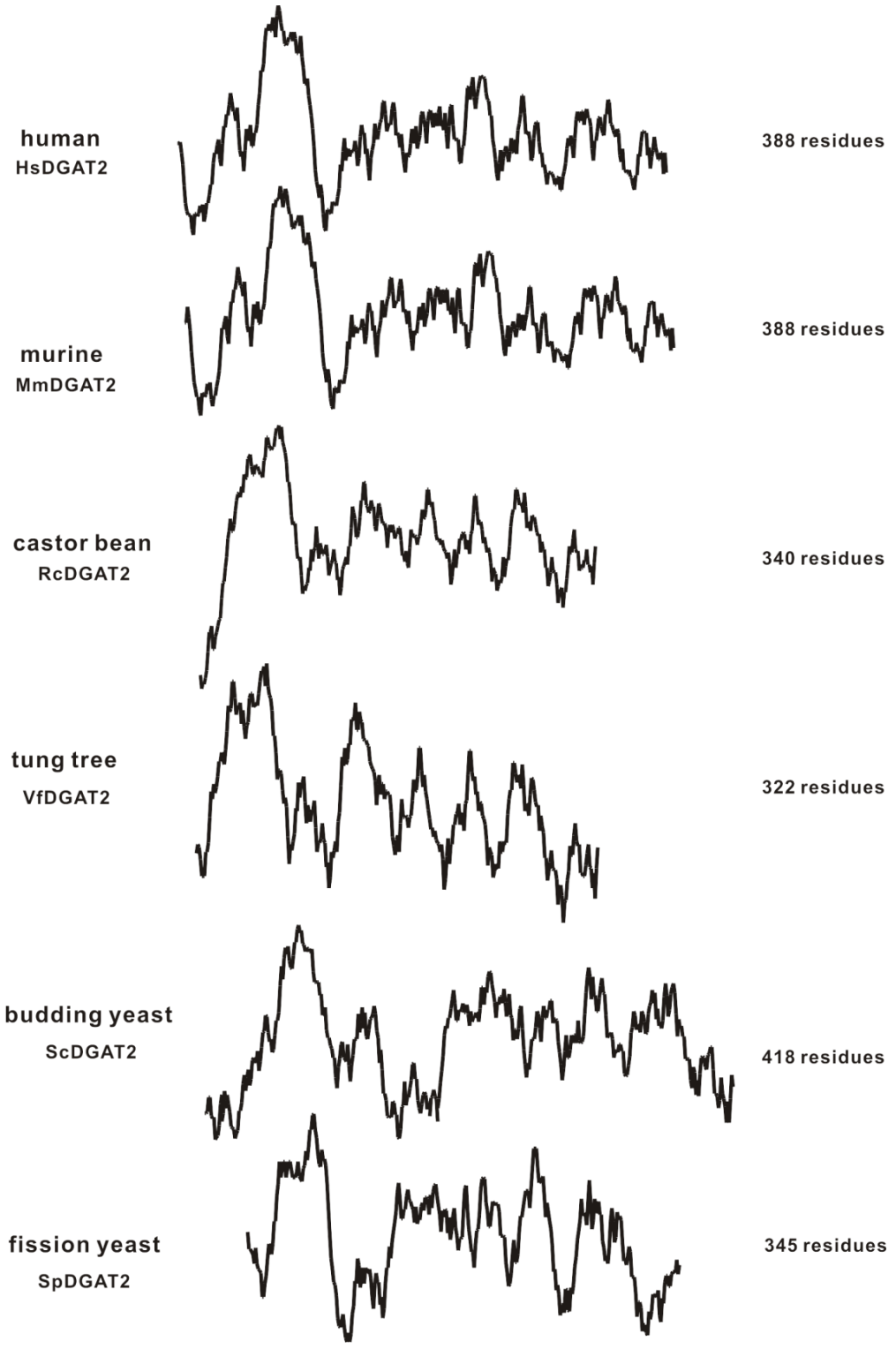
experimentally proven to reside in lumen, which was not the case in the model of murine DGAT2. This type of divergence in topologies between two protein homologues from different categories of organisms has been reported for other membrane proteins such as DGAT1 and serine palmitoyltransferase (Han et al., 2004; McFie et al., 2010).

These structural differences could also be reflected in the unique functional features of ScDGAT2. For instance, the unique hydrophilic stretch in the ER luminal loop, absent in murine DGAT2, was shown to influence enzyme activity in a sequence-dependent fashion (Figure 5.4B and 5.4C). It would be worthwhile to mimic this insertion in other DGAT2s to further clarify the role of this unique sequence. Moreover, ScDGAT2 lacks the consensus FLXLXXXn proposed to be the putative lipid-binding site in murine DGAT2, showing sequence ^{71}FVLFSIF^{77} instead. Thus, it is not surprising that point mutation on F71 ad L73 did not cause significant loss of the enzyme activity. Removal of the first TMD including these residues abolished activity (Figure 5.4A), indicating that a refinement of this binding domain is needed for yeast DGAT2.

Some of the functional and structural features of DGAT2 discussed here might have resulted from evolutionary adjustments specific to mammals, fungi and plants. Indeed, site-directed mutagenesis of HPHG to the plant specific

EPHS motif resulted in complete loss of ScDGAT2 activity, corroborating this hypothesis (Figure 5.4). Moreover, attempts to recover the activity of nonfunctional *Arabidopsis* DGAT2 by replacing the EPHS motif with the mammalian/fungal HPHG motif were unsuccessful (Chapter3; unpublished data). Furthermore, hydropathy analysis of functionally known DGAT2s presents divergences in the profiles between DGAT2s from animals, plants and fungi (Figure 5.12).

Figure 5.12. Hydropathy plot profiles of representative DGAT2s from animals, plants and fungi which have been functionally characterized. The Kyte-Doolittle plots were generated as described for Figure 5.3. Abbreviations for DGAT2 protein indicating the organism of origin and corresponding accession numbers are indicated under "Materials and methods."


human
HsDGAT2
388 residues
murine
MmDGAT2
388 residues
castor bean
RcDGAT2
340 residues
tung tree
VfDGAT2
322 residues
budding yeast
ScDGAT2
418 residues
fission yeast
SpDGAT2
345 residues

In conclusion, we have provided new insights into the functional and structural characteristics of DGAT2 from *S. cerevisiae* using a combination of experimental and computational analysis. An updated topology model for ScDGAT2 has been proposed, which can serve as a basis for further structure-function studies on this enzyme. The library of single cysteine residue-containing mutants could be used for more detailed structural analysis using biophysical approaches such as NMR spectroscopy to yield structural information with high-resolution (Van Horn et al., 2009), which is indispensable for validating the topology model and gaining of better understanding of the enzyme function.

6. General Discussion

Acyl-CoA:diacylglycerol acyltransferase 2 (DGAT2) has been extensively studied because of its important but diverse roles in triacylglycerol (TAG or TG) metabolism in different organisms. There is little structural and functional information available for this class of enzyme and its catalytic mechanism is not yet understood. Thus, the goal of this thesis was to generate novel information about the functional and structural characteristics of DGAT2 with an emphasis on signature residues and/or motifs in the protein.

The first part of this work focused on selecting a functional enzyme candidate as a model protein for further structure/function studies. Of all the tested *DGAT2* cDNAs from plants or yeast, only *DGAT2* from *Saccharomyces cerevisiae* (*ScDGAT2*) could be functionally expressed in the yeast expression system. *In vitro* and *in vivo* enzyme activities of other plant DGAT2 orthologues, including *Arabidopsis thaliana* and *Brassica napus DGAT2*s, were undetectable. Transcript and polypeptide analysis indicated that expression of plant *DGAT2*s under evaluation was blocked at the translational level, possibly because of the lack of unknown *in vivo* factors present in plants or differences in codon usage bias between yeast and plant species. My results from these plant *DGAT2*s are consistent with findings in recent reports based on investigations with *A. thaliana*. Although transcriptional expression of *AtDGAT2* was found in seeds and other

tissues, the T-DNA insertion mutant of *DGAT2* did not show any decrease of the oil content compared with the wild type (Li et al., 2010; Zhang et al., 2009). Moreover, a double mutant of *DGAT1* and *DGAT2* did not have substantial additional decrease in oil content in comparison with that already observed for the *DGAT1* single mutant (Zhang et al., 2009). Thus, it is possible that in *A. thaliana* and *B. napus*, DGAT2 does not play a significant role in TAG synthesis in seeds. Therefore, ScDGAT2 was chosen as the target protein for subsequent studies.

The second part of my research was aimed at understanding the role of seven cysteine residues in ScDGAT2. DGAT activities from the microsomal preparations were previously reported to be sensitive to cysteine-specific reagents (Kamisaka et al., 1993; Lozeman et al., 2001; Sauro and Strickland, 1990). ScDGAT2 was susceptible to thiol-modification as verified by various thiol-specific reagents including iodoacetamide (IA), N-ethylmaleimide (NEM) and 5, 5'-dithiobis-2-nitrobenzoate (DTNB), suggesting that cysteine residues may play a functional role in ScDGAT2. Site-directed mutagenesis (SDM) of cysteine residues, however, did not substantially affect the enzyme activity or protein expression levels. Also, no disulfide linkages between cysteine residues were identified. The enzyme inhibition mediated by NEM was localized to Cys^{314}, which is in close proximity to a highly conserved motif in DGAT2s from a wide

range of organisms. Thus, although none of the cysteine residues have a direct catalytic or structural role in ScDGAT2 function, Cys^{314} may be near a putative active site in this functional motif or may be involved in protein folding in a way not related to disulfide bonding.

The last part of the thesis research investigated various signature motifs in ScDGAT2 and their topological orientation. Based on *in silico* analysis, the functional significance and topological distribution of signature motifs in ScDGAT2 were determined using chemical modification in combination with SDM similarly to what was described early in thesis where the role of cysteine residues was probed. Both termini of the protein were demonstrated to be in the cytosol, but deletion mutagenesis suggested that the C-terminus may play a more essential role in maintaining enzyme activity. A hydrophilic stretch, unique for ScDGAT2, is located in a long ER luminal loop following the first transmembrane domain (TMD) was shown to influence enzyme activity in a sequence-dependent fashion. A highly conserved motif $^{129}YFP^{131}$ preceding this exclusive stretch resides in the same ER luminal loop, and was demonstrated to be important for enzyme catalysis. Another widely conserved residue, H^{195}, within the motif $^{193}HPHG^{196}$ could contribute to a putative active site possibly embedded in the membrane bilayer. Along with *in silico* analysis, topology studies

conducted using mutagenesis and chemical modification indicated a four-TMD model for ScDGAT2. This model shows some similarities to the previously proposed model form murine DGAT2 (Stone et al., 2006), but also reveals prominent differences, suggesting topological divergence of different DGAT2s.

Investigation of the functional importance of signature residue and/or motifs in ScDGAT2, revealed several strongly conserved motifs which are essential for ScDGAT2 catalysis including ^{129}YFP^{131}, the hydrophilic stretch exclusive to ScDGAT2 (amino acid residues 150 to 187) and ^{288}(RXGFX(K/R)XAXXXGXX(L/V)VPXXXFG(E/Q)^{311}. It will be worthwhile to further identify the irreplaceable residues in these motifs by site-saturation mutagenesis combined with directed evolution as we have employed for DGAT1 studies (Siloto et al., 2009b; 2009c). This approach allows the substitution of specific amino acid residues in the target polypeptide against all 20 possible amino acids at once. The activity of the recombinant enzyme mutants can be quickly quantified using a yeast live-cell-based fluorescence assay using Nile red dye specific for neutral lipids. These studies may provide us with further insight into the mechanism of DGAT action. Site-saturation mutagenesis and directed evolution may also lead to the generation of DGAT2 mutants with increased

catalytic efficiency and/or modulated substrate selectivity, which could be used in oilseed engineering for industrial applications

The combination of chemical modification and SDM offers many benefits for probing structure/function in membrane proteins such as DGAT2. Firstly, these approaches are suitable for using microsomes, which far less time-consuming and challenging to prepare compared to purified proteins. Moreover, these methodologies could be used for determining both the functional and structural role of signature residues and/or motifs as demonstrated in Chapter 4 and 5. When used for determining the topological organization of membrane proteins, chemical modification and SDM have several advantages over the other commonly used epitope tagging combined with immunofluorescence or protease protection assay for eukaryotic membrane proteins (van Geest and Lolkema, 2000). Introduction of a single cysteine residue in many positions is tolerated since the cysteine is non-bulky. In addition, this modification analysis is performed with complete, active and abundantly expressed membrane-incorporated protein mutant(s), which are likely to retain similar structure to the wild type protein. Additionally, the generated library of functional cysteine mutants could be used as a starting point for further biophysical studies. Major insights into structure/function in DGAT2 will come about if the protein is

eventually purified and effectively reconstituted into a membrane environment. This type of enzyme preparation would be amenable to structural analysis using biophysical approaches such as NMR spectroscopy.

There is an inherent limitation of using cysteine residue modification in mapping the topology of membrane proteins. When an introduced cysteine residue is in a highly protected region of the protein or is located at the interface of a protein-protein interaction, it might have limited access to thiol modification (Bogdanov et al., 2005; van Geest and Lolkema, 2000). Under this condition, assignment of a membrane location to this cysteine residue needs to proceed with caution. Analysis of the hydrophobicity of the segment can facilitate interpretation of the results but exhaustive scanning at nearby positions may be necessary as in the case of the Cys^{127} in ScDGAT2 (to determine the orientation of motif $^{129}YFP^{131}$, discussed in Chapter 5). The sidedness of this introduced cysteine residue was deduced in the ER lumen only after two cysteine residues at flanking positions (Cys^{114} and Cys^{183}) were proven to be in the ER lumen as well. A possible strategy to circumvent this hurdle is to use more complicated combinations of sulfhydryl reagents that differ in size and hydrophilicity, as demonstrated in the investigation of topological organization of acyl-CoA:cholesterol acyltransfease 1 (ACAT1) expressed in mammalian cells (Guo et

al., 2005b). With the further development of this strategy in yeast cells, the proposed four-TMD model could be confirmed and/or improved by verifying the orientation of regions in the model established only by computational constraints. In addition, as indicated previously, the different topological characteristics between yeast and murine DGAT2s might have resulted from evolutionary adaptation. Furthermore, it will be interesting to explore the membrane topology of DGAT2 orthologues from plants to further test this hypothesis.

In conclusion, a combination of SDM and chemical modification were used to probe structure/function in a DGAT2. We await the discovery of three-dimensional structure of DGAT2s, which will validate the findings described here and allow us to deeply and comprehensively understand the molecular mechanism of this enzyme.

7. Reference List

Abell, B. M., M. Hahn, L. A. Holbrook & M. M. Moloney (2004) Membrane topology and sequence requirements for oil body targeting of oleosin. *Plant Journal,* 37**,** 461-470.

Alvarez, H. M. & A. Steinbuchel (2002) Triacylglycerols in prokaryotic microorganisms. *Applied Microbiology and Biotechnology,* 60**,** 367-376.

Angov, E., C. J. Hillier, R. L. Kincaid, and J. A. Lyon (2008) Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host: *Plos One***,** 3.

Arai, M., H. Mitsuke, M. Ikeda, J. X. Xia, T. Kikuchi, M. Satake & T. Shimizu (2004) ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Research,* 32**,** W390-W393.

Athenstaedt, K. & G. Daum (2006) The life cycle of neutral lipids: synthesis, storage and degradation. *Cellular and Molecular Life Sciences,* 63**,** 1355-1369.

Bagnato, C. & R. A. Igal (2003) Overexpression of diacylglycerol acyltransferase-1 reduces phospholipid synthesis, proliferation, and invasiveness in simian virus 40-transformed human lung fibroblasts. *Journal of Biological Chemistry,* 278**,** 52203-52211.

Bogdanov, M., W. Zhang, J. Xie & W. Dowhan (2005) Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM^{TM}): Application to lipid-specific membrane protein topogenesis. *Methods,* 36**,** 148-171.

Bouvier-Nave, P., Benveniste, P., Noiriel, A., and Schaller, H. (2000a). Expression in yeast of an acyl-CoA:diacylglycerol acyltransferase cDNA

from *Caenorhabditis elegans*. *Biochemical Society Transactions,* 28**,** 692-695.

Bouvier-Nave, P., Benveniste, P., Oelkers, P., Sturley, S.L., and Schaller, H. (2000b). Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *European Journal of Biochemistry*, 267**,** 85-96.

Boyd, D. 1996. Use of gene fusions to determine membrane protein topology. In *Membrane protein structure. Experimental approaches,* ed. S. H. White, 144-163. New york: Oxford University Press.

Burgal, J., J. Shockey, C. F. Lu, J. Dyer, T. Larson, I. Graham & J. Browse (2008) Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnology Journal,* 6**,** 819-831.

Cahoon, E. B., J. M. Shockey, C. R. Dietrich, S. K. Gidda, R. T. Mullen & J. M. Dyer (2007) Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr.Opin.Plant Biol.,* 10**,** 236-244.

Cases, S., S. J. Smith, Y. W. Zheng, H. M. Myers, S. R. Lear, E. Sande, S. Novak, C. Collins, C. B. Welch, A. J. Lusis, S. K. Erickson & R. V. Farese. 1998. Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. 13018-13023.

Cases, S., S. J. Stone, P. Zhou, E. Yen, B. Tow, K. D. Lardizabal, T. Voelker & R. V. Farese (2001) Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. *Journal of Biological Chemistry,* 276**,** 38870-38876.

Chang, C. C. Y., H. Y. Huh, K. M. Cadigan, and T. Y. Chang (1993) Molecular-Cloning and functional expression of human acyl-Coenzyme-A cholesterol acyltransferase cDNA in mutant Chinese-hamster ovary cells: *Journal of Biological Chemistry,* 268**,** 20747-20755.

Chen, H. C. & R. V. Farese (2000) DGAT and triglyceride synthesis: A new target for obesity treatment? *Trends in Cardiovascular Medicine,* 10**,** 188-192.

Chen, H. C., Z. Ladha & R. V. Farese (2002a) Deficiency of acyl coenzyme A : diacylglycerol acyltransferase 1 increases leptin sensitivity in murine obesity models. *Endocrinology,* 143**,** 2893-2898.

Chen, H. C., S. J. Smith, Z. Ladha, D. R. Jensen, L. D. Ferreira, L. K. Pulawa, J. G. McGuire, R. E. Pitas, R. H. Eckel & R. V. Farese (2002b) Increased insulin and leptin sensitivity in mice lacking acyl CoA : diacylglycerol acyltransferase 1. *Journal of Clinical Investigation,* 109**,** 1049-1055.

Cheng, D., T. C. Nelson, J. Chen, S. G. Walker, J. Wardwell-Swanson, R. Meegalla, R. Taub, J. T. Billheimer, M. Ramaker & J. N. Feder (2003) Identification of acyl coenzyme A : monoacylglycerol acyltransferase 3, an intestinal specific enzyme implicated in dietary fat absorption. *Journal of Biological Chemistry,* 278**,** 13611-13614.

Chiapello, H., F. Lisacek, M. Caboche, and A. Henaut (1998) Codon usage and gene function are related in sequences of *Arabidopsis thaliana*. *Gene*, 209**,** GC1-GC38.

Choi, C. S., D. B. Savage, A. Kulkarni, X. X. Yu, Z. X. Liu, K. Morino, S. Kim, A. Distefano, V. T. Samuel, S. Neschen, D. Y. Zhang, A. Wang, X. M. Zhang, M. Kahn, G. W. Cline, S. K. Pandey, J. G. Geisler, S. Bhanot, B. P. Monia & G. I. Shulman (2007) Suppression of diacylglycerol acyltransferase-2 (DGAT2), but not DGAT1, with antisense oligonucleotides reverses diet-induced hepatic steatosis and insulin resistance. *Journal of Biological Chemistry,* 282**,** 22678-22688.

Claros, M. G. & G. Vonheijne (1994) Toppred-II-an improved software for membrane-protein structure predictions. *Computer Applications in the Biosciences,* 10**,** 685-686.

Dahlqvist, A., U. Stahl, M. Lenman, A. Banas, M. Lee, L. Sandager, H. Ronne, and H. Stymne (2000) Phospholipid : diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of

triacylglycerol in yeast and plants. *Proceedings of the National Academy of Sciences of the United States of America*, 97**,** 6487-6492.

Dobson, C. M. (2003) Protein folding and misfolding. *Nature,* 426**,** 884-890.

Dobson, C. M., A. Sali & M. Karplus (1998) Protein folding: A perspective from theory and experiment. *Angewandte Chemie-International Edition,* 37**,** 868-893.

Farese, R. V., S. Cases & S. J. Smith (2000) Triglyceride synthesis: insights from the cloning of diacylglycerol acyltransferase. *Current Opinion in Lipidology,* 11**,** 229-234.

Ferrer, J. L., J. M. Jez, M. E. Bowman, R. A. Dixon & J. P. Noel (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nature Structural Biology,* 6**,** 775-784.

Frentzen, M. (1998) Acyltransferases from basic science to modified seed oils. *Fett-Lipid,* 100**,** 161-166.

Gangar, A., A. A. Karande, and R. Rajasekharan (2001) Isolation and localization of a cytosolic 10 S triacylglycerol biosynthetic multienzyme complex from oleaginous yeast. *Journal of Biological Chemistry,* 276**,** 10290-10298.

Gao, F. P. & T. A. Cross (2005) Recent developments in membrane-protein structural genomics. *Genome Biology,* 6.

Gietz, R. D. & R. H. Schiestl (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols,* 2**,** 31-34.

Gilstring, C. F. & P. O. Ljungdahl (2000) A method for determining the *in vivo* topology of yeast polytopic membrane proteins demonstrates that Gap1p fully integrates into the membrane independently of Shr3p. *Journal of Biological Chemistry,* 275**,** 31488-31495.

Graham, I. A. (2008) Seed storage oil mobilization. *Annual Review of Plant Biology,* 59**,** 115-142.

Guo, Z. Y., C. C. Y. Chang, X. H. Lu, J. Chen, B. L. Li & T. Y. Chang (2005a) The disulfide linkage and the free sulfhydryl accessibility of acyl-coenzyme A : Cholesterol acyltransferase 1 as studied by using mPEG5000-maleimide. *Biochemistry,* 44**,** 6537-6546.

Guo, Z. Y., S. Lin, J. A. Heinen, C. C. Y. Chang & T. Y. Chang (2005b) The active site His-460 of human acyl-coenzyme A : cholesterol acyltransferase 1 resides in a hitherto undisclosed transmembrane domain. *Journal of Biological Chemistry,* 280**,** 37814-37826.

Han, G. S., K. Gable, L. Y. Yan, M. Natarajan, J. Krishnamurthy, S. D. Gupta, A. Borovitskaya, J. M. Harmon & T. M. Dunn (2004) The topology of the Lcb1p subunit of yeast serine palmitoyltransferase. *Journal of Biological Chemistry,* 279**,** 53707-53716.

Hara, A. & N. S. Radin (1978) Lipid extractio of tissues with a low-toxicity solvent. *Analytical Biochemistry,* 90**,** 420-426.

Heckman, K. L. & L. R. Pease (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols,* 2**,** 924-932.

Herman, E. M. (2008) Endoplasmic reticulum bodies: solving the insoluble. *Current Opinion in Plant Biology,* 11**,** 672-679.

Hirokawa, T., S. Boon-Chieng & S. Mitaku (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics,* 14**,** 378-379.

Hishikawa, D., H. Shindou, S. Kobayashi, H. Nakanishi, R. Taguchi & T. Shimizu (2008) Discovery essential of a lysophospholipid acyltransferase family for membrane asymmetry and diversity. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 2830-2835.

Hobbs, D. H. & M. J. Hills (2000) Expression and characterization of diacylglycerol acyltransferase from *Arabidopsis thaliana* in insect cell cultures. *Biochemical Society Transactions,* 28**,** 687-689.

Hobbs, D. H., C. F. Lu & M. J. Hills (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. *Febs Letters,* 452**,** 145-149.

Hofmann, K. (2000) A superfamily of membrane-bound O-acyltransferases with implications for Wnt signaling. *Trends in Biochemical Sciences,* 25**,** 111-112.

Hofmann, K. & W. Stoffel (1993) TM base-A database of membrane spanning proteins segments. *Biol. Chem. Hopper-Seyler,* 374**,** 166.

Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman & E. K. O'Shea (2003) Global analysis of protein localization in budding yeast. *Nature,* 425**,** 686-691.

Inokoshi, J., Kawamoto, K., Takagi, Y., Matsuhama, M., Omura, S., and Tomoda, H. (2009). Expression of two human acyl-CoA:diacylglycerol acyltransferase isozymes in yeast and selectivity of microbial inhibitors toward the isozymes. *Journal of Antibiotics*, 62**,** 51-54

Jako, C., A. Kumar, Y. D. Wei, J. T. Zou, D. L. Barton, E. M. Giblin, P. S. Covello & D. C. Taylor (2001) Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiology,* 126**,** 861-874.

Kalscheuer, R., and A. Steinbuchel (2003) A novel bifunctional wax ester synthase/acyl-CoA : diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1. *Journal of Biological Chemistry,* 278**,** 8075-8082.

Kamisaka & Yasushi (2010) Activation of diacylglycerol acyltransferase expressed in *Saccharomyces cerevisiae*: overexpression of Dga1p lacking the N-terminal region in the Δsnf2 disruptant produces a significant increase in its enzyme activity. *Applied Microbiology and Biotechnology,* 88.

Kamisaka, Y., T. Yokochi, T. Nakahara & O. Suzuki (1993) Characterization of the diacylglycerol acyltransferase activity in the membrane-fraction from a fungus. *Lipids,* 28**,** 583-587.

Karantonis, H. C., T. Nomikos & C. A. Demopoulos (2009) Triacylglycerol metabolism. *Current Drug Targets,* 10**,** 302-319.

Katavic, V., D. W. Reed, D. C. Taylor, E. M. Giblin, D. L. Barton, J. T. Zou, S. L. MacKenzie, P. S. Covello & L. Kunst (1995) Alteration of seed fatty-acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. *Plant Physiology,* 108**,** 399-409.

Kaup, M. T., C. D. Froese & J. E. Thompson (2002) A role for diacylglycerol acyltransferase during leaf senescence. *Plant Physiology,* 129**,** 1616-1626.

King, A. J., J. A. Segreti, K. J. Larson, A. J. Souers, P. R. Kym, R. M. Reilly, G. Zhao, S. W. Mittelstadt & B. F. Cox (2009) Diacylglycerol acyltransferase 1 inhibition lowers serum triglycerides in the zucker fatty rat and the hyperlipidemic hamster. *Journal of Pharmacology and Experimental Therapeutics,* 330**,** 526-531.

Krogh, A., B. Larsson, G. von Heijne & E. L. L. Sonnhammer (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology,* 305**,** 567-580.

Kroon, J. T. M., W. X. Wei, W. J. Simon & A. R. Slabas (2006) Identification and functional expression of a type 2 acyl-CoA : diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. *Phytochemistry,* 67**,** 2541-2549.

Kuerschner, L., C. Moessinger & C. Thiele (2008) Imaging of lipid biosynthesis: how a neutral lipid enters lipid droplets. *Traffic,* 9**,** 338-352.

Kyte, J. & R. F. Doolittle (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* 157**,** 105-132.

Lardizabal, K., R. Effertz, C. Levering, J. Mai, M. C. Pedroso, T. Jury, E. Aasen, K. Gruys & K. Bennett (2008) Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiology,* 148**,** 89-96.

Lardizabal, K. D., J. T. Mai, N. W. Wagner, A. Wyrick, T. Voelker & D. J. Hawkins (2001) DGAT2 is a new diacylglycerol acyltransferase gene family - purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. *Journal of Biological Chemistry,* 276**,** 38862-38869.

Li, F., X. Wu, P. Lam, D. Bird, H. Zheng, L. Samuels, R. Jetter, and L. Kunst (2008) Identification of the wax ester synthase/acyl-coenzyme A: Diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in *Arabidopsis*. *Plant Physiology,* 148**,** 97-107.

Li, R. Z., K. S. Yu & D. F. Hildebrand (2010) DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. *Lipids,* 45**,** 145-157.

Liang, J.J., Oelkers, P., Guo, C.Y., Chu, P.C., Dixon, J.L., Ginsberg, H.N., and Sturley, S.L. (2004). Overexpression of human diacylglycerol acyltransferase 1, acyl-CoA:cholesterol acyltransferase 1, or acyl-CoA : cholesterol acyltransferase 2 stimulates secretion of apolipoprotein B-containing lipoproteins in McA-RH7777 cells. *Journal of Biological Chemistry*, 279**,** 44938-44944.

Little, D., R. Weselake, K. Pomeroy, T. Furukawastoffer & J. Bagu (1994) Solubilization and characterization of diacylglycerol acyltransferase from microspore-derived cultures of oilseed rape. *Biochemical Journal,* 304**,** 951-958.

Liu, Q., R. M. P. Siloto & R. J. Weselake (2010) Role of cysteine residues in thiol modification of acyl-CoA:diacylglycerol acyltransferase 2 from yeast. *Biochemistry,* 49**,** 3237-3245.

Liu, Y., J. S. Millar, D. A. Cromley, M. Graham, R. Crooke, J. T. Billheimer & D. J. Rader (2008) Knockdown of acyl-CoA : diacylglycerol acyltransferase 2 with antisense oligonucleotide reduces VLDL TG and ApoB secretion in mice. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids,* 1781**,** 97-104.

Lozeman, F. J., C. K. Middleton, J. T. Deng, E. C. Kazala, C. Verhaege, P. S. Mir, A. Laroche, D. R. C. Bailey & R. J. Weselake (2001) Characterization of microsomal diacylglycerol acyltransferase activity from bovine adipose and muscle tissue. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* 130**,** 105-115.

Lu, C. F. L., S. B. de Noyer, D. H. Hobbs, J. L. Kang, Y. C. Wen, D. Krachtus & M. J. Hills (2003) Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in triacylglycerol biosynthesis, in Arabidopsis thaliana. *Plant Molecular Biology,* 52**,** 31-41.

Lu, J. L. & C. Deutsch (2001) Pegylation: A method for assessing topological accessibilities in Kv1.3. *Biochemistry,* 40**,** 13288-13301.

Lu, X. H., S. Lin, C. C. Y. Chang & T. Y. Chang (2002) Mutant acyl-coenzyme A : cholesterol acyltransferase 1 devoid of cysteine residues remains catalytically active. *Journal of Biological Chemistry,* 277**,** 711-718.

Lundstrom, K. (2004) Structural genomics on membrane proteins: mini review. *Combinatorial Chemistry & High Throughput Screening,* 7**,** 431-439.

Lung, S. C. & R. J. Weselake (2006) Diacylglycerol acyltransferase: A key mediator of plant triacylglycerol synthesis. *Lipids,* 41**,** 1073-1088.

Man, W. C., M. Miyazaki, K. Chu & J. Ntambi (2006) Colocalization of SCD1 and DGAT2: implying preference for endogenous monounsaturated fatty acids in triglyceride synthesis. *Journal of Lipid Research,* 47**,** 1928-1939.

Marmorstein, R. (2001) Structure and function of histone acetyltransferases. *Cellular and Molecular Life Sciences,* 58**,** 693-703.

Mavraganis, I., D. Meesapyodsuk, P. Vrinten, M. Smith & X. Qiu (2010) Type II diacylglycerol acyltransferase from *Claviceps purpurea* with ricinoleic acid, a hydroxyl fatty acid of industrial importance, as preferred substrate. *Applied and Environmental Microbiology*, 76**,** 1135-1142.

McFie, P. J., S. L. Stone, S. L. Banman & S. J. Stone (2010) Acyl coA:diacylglycerol acyltransferase-1 (DGAT1): topological orientation, identification of a putative active site histidine and the role of the N-terminus in dimer/tetramer formationn. *Journal of biological chemistry*. 10.1074/jbc.M110163691

McLuskey, K., A. W. Roszak, Y. S. Zhu & N. W. Isaacs (2010) Crystal structures of all-alpha type membrane proteins. *European Biophysics Journal with Biophysics Letters*, 39**,** 723-755.

Milcamps, A., Tumaney, A.W., Paddock, T., Pan, D.A., Ohlrogge, J., and Pollard, M. (2005). Isolation of a gene encoding a 1,2-diacylglycerol-sn-acetyl-CoA acetyltransferase from developing seeds of Euonymus alatus. *Journal of Biological Chemistry*, 280**,** 5370-5377.

Millar, J. S. & J. T. Billheimer. 2006. *Acyl CoA:Diacylglycerol acyltransferase (DGATs) as therapeutic targets for cardiovascular disease*. Oxon: Taylor & Francis.

Millar, J. S., S. J. Stone, U. J. F. Tietge, B. Tow, J. T. Billheimer, J. S. Wong, R. L. Hamilton, R. V. Farese & D. J. Rader (2006) Short-term overexpression of DGAT1 or DGAT2 increases hepatic triglyceride but not VLDL triglyceride or apoB production. *Journal of Lipid Research*, 47**,** 2297-2305.

Mulhall, B. P., J. P. Ong & Z. M. Younossi (2002) Non-alcoholic fatty liver disease: An overview. *Journal of Gastroenterology and Hepatology*, 17**,** 1136-1143.

Natter, K., P. Leitner, A. Faschinger, H. Wolinski, S. McCraith, S. Fields & S. D. Kohlwein (2005) The spatial organization of lipid synthesis in the yeast *Saccharomyces cerevisiae* derived from large scale green fluorescent

protein tagging and high resolution microscopy. *Molecular & Cellular Proteomics,* 4**,** 662-672.

Normington, K., K. Kohno, Y. Kozutsumi, M. J. Gething & J. Sambrook (1989) *S. cerevisiae* encodes an essential protein homologous in sequence and function to mammalian BIP. *Cell,* 57**,** 1223-1236.

Oelkers, P., D. Cromley, M. Padamsee, J. T. Billheimer & S. L. Sturley (2002) The DGA1 gene determines a second triglyceride synthetic pathway in yeast. *Journal of Biological Chemistry*, 277**,** 8877-8881.

Quittnat, F., Nishikawa, Y., Stedman, T.T., Voelker, D.R., Choi, J.Y., Zahn, M.M., Murphy, R.C., Barkley, R.M., Pypaert, M., Joiner, K.A., and Coppens, I. (2004). On the biogenesis of lipid bodies in ancient eukaryotes: synthesis of triacylglycerols by a Toxoplasma DGAT1-related enzyme. *Molecular and Biochemical Parasitology*, 138**,** 107-122.

Perry, H. J., R. Bligny, E. Gout & J. L. Harwood (1999) Changes in Kennedy pathway intermediates associated with increased triacylglycerol synthesis in oil-seed rape. *Phytochemistry,* 52**,** 799-804.

Persson, B. 2006. Bioinformatics in Membrane Protein Analysis. In *Structual genomics of membrane proteins,* ed. K. H. Lundstrom, 5-20. Florida: CRC press.

Rani, S. H., T. H. A. Krishna, S. Saha, A. S. Negi, and R. Rajasekharan (2010) Defective in cuticular ridges of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. *Journal of Biological Chemistry*. 10.1074/jbc.M110.133116.

Reinhart, C. & C. Krettler. 2006. Expression of membrane proteins in yeast. In *Structural genomics on membrane proteins,* ed. K. Lundstorm, 115-152. Boca Raton: CRC press.

Romanos, M. A., C. A. Scorer & J. J. Clare (1992) Foreign gene-expression in yeast-a review. *Yeast,* 8**,** 423-488.

Routaboul, J. M., C. Benning, N. Bechtold, M. Caboche & L. Lepiniec (1999) The TAG1 locus of *Arabidopsis* encodes for a diacylglycerol acyltransferase. *Plant Physiology and Biochemistry,* 37**,** 831-840.

Saha, S., B. Enugutti, S. Rajakumari, and R. Rajasekharan (2006) Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiology*, 141**,** 1533-1543.

Sandager, L., M. H. Gustavsson, U. Stahl, A. Dahlqvist, E. Wiberg, A. Banas, M. Lenman, H. Ronne & S. Stymne (2002) Storage lipid synthesis is non-essential in yeast. *Journal of Biological Chemistry,* 277**,** 6478-6482.

Sauro, V. S. & K. P. Strickland (1990) Triacylglycerol synthesis and diacylglycerol acyltransferase activity during skeletal myogenesis. *Biochemistry and Cell Biology,* 68**,** 1393-1401.

Sharp, P. M., E. Cowe, D. G. Higgins, D. C. Shields, K. H. Wolfe, and F. Wright (1988) Codon usage patterns in *Escherichia-coli*, *Bacillus-subitlis*, *Saccharomyces-pombe*, *Drosophila-melanogaster* and *Homo-sapiens*-a review of the considerable within-species diversity. *Nucleic Acids Research*, 16**,** 8207-8211.

Sharp, P. M., T. M. F. Tuohy, and K. R. Mosurski (1986) Codon usage in yeast-cluster-analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Research*,14**,** 5125-5143.

Shockey, J. M., S. K. Gidda, D. C. Chapital, J. C. Kuan, P. K. Dhanoa, J. M. Bland, S. J. Rothstein, R. T. Mullen & J. M. Dyer (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell,* 18**,** 2294-2313.

Siloto, R. M. P., Liu, Q., Weselake R. J., He, X. H. & McKeon, T. 2009a. Insights into structure and function of acyl-CoA:diacylglycerol acyltransferase. In *Biocatalysis and Molecular Engineering,* ed. C. Hou. John Wiley & Sons.

Siloto, R. M. P., M. Truksa, D. Brownfield, A. G. Good & R. J. Weselake (2009b) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiology and Biochemistry,* 47**,** 456-461.

Siloto, R. M. P., M. Truksa, X. H. He, T. McKeon & R. J. Weselake (2009c) Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system. *Lipids,* 44**,** 963-973.

Smith, S. J., S. Cases, D. R. Jensen, H. C. Chen, E. Sande, B. Tow, D. A. Sanan, J. Raber, R. H. Eckel & R. V. Farese (2000) Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. *Nature Genetics,* 25**,** 87-90.

Snyder, C. L., Yurchenko, O. P., Siloto, R. M. P., Chen, X., Liu, Q., Mietkiewska, E. & Weselake, R. J. (2009) Acyltransferase action in the modification of seed oil biosynthesis. *New Biotechnology,* 26**,** 11-16.

Sorger, D. & G. Daum (2002) Synthesis of triacylglycerols by the acyl-coenzyme A: Diacyl-glycerol acyltransferase Dga1p in lipid particles of the yeast *Saccharomyces cerevisiae*. *Journal of Bacteriology,* 184**,** 519-524.

Sorger, D., and G. Daum (2003) Triacylglycerol biosynthesis in yeast. *Applied Microbiology and Biotechnology,* 61**,** 289-299.

Stals, H. K., W. Top & P. E. Declercq (1994) Regulation of triacylglycerol synthesis in permeabilized rat hepatocytes-role of fatty-acid concentration and diacylglycerol acyltransferase. *Febs Letters,* 343**,** 99-102.

Stone, S. J., M. C. Levin & R. V. Farese (2006) Membrane topology and identification of key functional amino acid residues of murine acyl-CoA: diacylglycerol acyltransferase-2. *Journal of Biological Chemistry,* 281**,** 40273-40282.

Stone, S. J., M. C. Levin, P. Zhou, J. Y. Han, T. C. Walther & R. V. Farese (2009) The Endoplasmic reticulum enzyme DGAT2 is found in mitochondria-associated membranes and has a mitochondrial targeting signal that

promotes its association with mitochondria. *Journal of Biological Chemistry,* 284**,** 5352-5361.

Stone, S. J., H. M. Myers, S. M. Watkins, B. E. Brown, K. R. Feingold, P. M. Elias & R. V. Farese (2004) Lipopenia and skin barrier abnormalities in DGAT2-deficient mice. *Journal of Biological Chemistry,* 279**,** 11767-11776.

Stoveken, T., R. Kalscheuer, U. Malkus, R. Reichelt, and A. Steinbuchel (2005) The wax ester synthase/acyl coenzyme A : diacylglycerol acyltransferase from *Acinetobacter sp* strain ADP1: Characterization of a novel type of acyltransferase. *Journal of Bacteriology,* 187**,** 1369-1376.

Tamura, N., S. Konishi, S. Iwaki, T. Kimura-Someya, S. Nada & A. Yamaguchi (2001) Complete cysteine-scanning mutagenesis and site-directed chemical modification of the Tn10-encoded metal-tetracycline/H+ antiporter. *Journal of Biological Chemistry,* 276**,** 20330-20339.

Taylor, D., C., Zhang. Y., Kumar., A., Francis, T., Giblin, M., E., Barton, D. L., Ferrie, J. R., Laroche, A., Shah, S., Zhu, W. M., Snyder, C. L., Hall, L. H., Rakow, G., Harwood, J. L. & Weselake, R. J. (2009). Molecular modification of triacylglycerol accumulation by over-expression of *DGAT1* to produce canola with increased seed oil content under field conditions. *Botany*. 87**,** 533-543.

Turkish, A. R., A. L. Henneberry, D. Cromley, M. Padamsee, P. Oelkers, H. Bazzi, A. M. Christiano, J. T. Billheimer & S. L. Sturley (2005) Identification of two novel human Acyl-CoA wax alcohol acyltransferases - Members of the diacylglycerol acyltransferase 2 (DGAT2) gene superfamily. *Journal of Biological Chemistry,* 280**,** 14755-14764.

Tusnady, G. E. & I. Simon (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics,* 17**,** 849-850.

Tzen, J. T. C., G. C. Lie & A. H. C. Huang (1992) Characterization of the charged components and their topology on the surface of plant seed oil bodies. *Journal of Biological Chemistry,* 267**,** 15626-15634.

van Geest, M. & J. S. Lolkema (2000) Membrane topology and insertion of membrane proteins: Search for topogenic signals. *Microbiology and Molecular Biology Reviews,* 64**,** 13-33.

Van Horn, W. D., H. J. Kim, C. D. Ellis, A. Hadziselimovic, E. S. Sulistijo, M. D. Karra, C. L. Tian, F. D. Sonnichsen & C. R. Sanders (2009) Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science,* 324**,** 1726-1729.

Wakimoto, K., H. Chiba, H. Michibata, M. Seishima, S. Kawasaki, K. Okubo, H. Mitsui, H. Torii & Y. Imai (2003) A novel diacylglycerol acyltransferase (DGAT2) is decreased in human psoriatic skin and increased in diabetic mice. *Biochemical and Biophysical Research Communications,* 310**,** 296-302.

Wang, Y., M. Toei & M. Forgac (2008) Analysis of the membrane topology of transmembrane segments in the C-terminal hydrophobic domain of the yeast vacuolar ATPase subunit a (Vph1p) by chemical modification. *Journal of Biological Chemistry,* 283**,** 20696-20702.

Weiss, S. B., and E. P. Kennedy (1956) The Enzymatic Synthesis of Triglycerides: *Journal of the American Chemical Society,* 78**,** 3550-3550.

Weselake, R. J. 2005. Storage Lipids. In *Plant Lipids: Biology, Utilization and Manipulation,* ed. M. D.J., 162-225. Oxford: Blackwell publishing.

Weselake, R. J., C. L. Nykiforuk, A. Laroche, N. A. Patterson, W. B. Wiehler, S. J. Szarka, M. M. Moloney, L. W. Tari & U. Derekh (2000) Expression and properties of diacylglycerol acyltransferase from cell-suspension cultures of oilseed rape. *Biochemical Society Transactions,* 28**,** 684-686.

Weselake, R. J., S. Shah, M. G. Tang, P. A. Quant, C. L. Snyder, T. L. Furukawa-Stoffer, W. M. Zhu, D. C. Taylor, J. T. Zou, A. Kumar, L. Hall, A. Laroche, G. Rakow, P. Raney, M. M. Moloney & J. L. Harwood (2008) Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. *Journal of Experimental Botany,* 59**,** 3543-3549.

Weselake, R. J., D. C. Taylor, M. K. Pomeroy, S. L. Lawson & E. W. Underhill (1991) Properties of diacylglycerol acyltransferase from microspore-derived embryos of *Brassica-napus*. *Phytochemistry,* 30**,** 3533-3538.

Woycechowsky, K. J. & R. T. Raines (2000) Native disulfide bond formation in proteins. *Current Opinion in Chemical Biology,* 4**,** 533-539.

Xu, J. Y., T. Francis, E. Mietkiewska, E. M. Giblin, D. L. Barton, Y. Zhang, M. Zhang & D. C. Taylor (2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (*DGAT1*) gene from Tropaeolum majus, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. *Plant Biotechnology Journal,* 6**,** 799-818.

Yen, C. L. E., C. H. Brown, M. Monetti & R. V. Farese (2005) A human skin multifunctional O-acyltransferase that catalyzes the synthesis of acylglycerols, waxes, and retinyl esters. *Journal of Lipid Research,* 46**,** 2388-2397.

Yen, C. L. E. & R. V. Farese (2003) MGAT2, a monoacylglycerol acyltransferase expressed in the small intestine. *Journal of Biological Chemistry,* 278**,** 18532-18537.

Yen, C. L. E., S. J. Stone, S. Cases, P. Zhou & R. V. Farese (2002) Identification of a gene encoding MGAT1, a monoacylglycerol acyltransferase. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 8512-8517.

Yen, C. L. E., S. J. Stone, S. Koliwad, C. Harris & R. V. Farese (2008) DGAT enzymes and triacylglycerol biosynthesis. *Journal of Lipid Research,* 49**,** 2283-2301.

Yu, X. X., S. F. Murray, S. K. Pandey, S. L. Booten, D. J. Bao, X. Z. Song, S. Kelly, S. Y. Chen, R. McKay, B. P. Monia & S. Bhanot (2005) Antisense oligonucleotide reduction of DGAT2 expression improves hepatic steatosis and hyperlipidemia in obese mice. *Hepatology,* 42**,** 362-371.

Yu, K., Li, R., Hatanaka, T., and Hildebrand, D. (2008). Cloning and functional analysis of two type 1 diacylglycerol acyltransferases from *Vernonia galamensis*. *Phytochemistry*，69**,** 1119-1127.

Yuan, Z., J. S. Mattick & R. D. Teasdale (2004) SVMtm: Support vector machines to predict transmembrane segments. *Journal of Computational Chemistry*, 25**,** 632-636.

Yurchenko, O. 2009. Role of cytosolic acyl-CoA binding protein in seed oil Biosynthesis. In *Department of Agricultural, Food and Nutritional Science*, 195. Edmonton: University of alberta.

Zhang, M., J. L. Fan, D. C. Taylor & J. B. Ohlrogge (2009) DGAT1 and PDAT1 Acyltransferases have overlapping functions in *Arabidopsis* triacylglycerol biosynthesis and are essential for normal pollen and seed development. *Plant Cell*, 21**,** 3885-3901.

Zhang, Q., H. K. Chieu, C. P. Low, S. C. Zhang, C. K. Heng & H. Y. Yang (2003) Schizosaccharomyces pombe cells deficient in triacylglycerols synthesis undergo apoptosis upon entry into the stationary phase. *Journal of Biological Chemistry*, 278**,** 47145-47155.

Zhao, G., A. J. Souers, M. Voorbach, H. D. Falls, B. Droz, S. Brodjian, Y. Y. Lau, R. R. Iyengar, J. Gao, A. S. Judd, S. H. Waaaw, M. M. Ravn, K. M. Engstrom, J. K. Lynch, M. M. Mulhern, J. Freeman, B. D. Dayton, X. J. Wang, N. Grihalde, D. Fry, D. W. A. Beno, K. C. Marsh, Z. Su, G. J. Diaz, C. A. Collins, H. Sham, R. M. Reilly, M. E. Brune & P. R. Kym (2008) Validation of diacyl glycerolacyltransferase I as a novel target for the treatment of obesity and dyslipidemia using a potent and selective small molecule inhibitor. *Journal of Medicinal Chemistry*, 51**,** 380-383.

Zheng, P., Allen, W.B., Roesler, K., Williams, M.E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D., Solawetz, W., Bhattramakki, D., Llaca, V., Deschamps, S., Zhong, G.Y., Tarczynski, M.C., and Shen, B. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nature Genetics*，40**,** 367-372.

Zou, J. T., Y. D. Wei, C. Jako, A. Kumar, G. Selvaraj & D. C. Taylor (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *Plant Journal,* 19**,** 645-653.

A version of Appendix.1 has been published. Siloto, R. M. P, Liu, Q, Weselake, R. J., He, X. H. and McKeon, T. 2009. In Biocatalysis and Molecular Engineering, ed. C.Hou. John Wiley & Sons.

Appendix. 1 Insights into the Structure and Function of Acyl-CoA:Diacylglycerol Acyltransferase

(Reprinted with permission of John Wiley& Sons, Ltd.)

I was mainly responsible for writing sections on "*Membrane topological organization of DGATs*" and "*Subcelluar localization of DGATs*" as well as discussion on site-directed mutagenesis.

Contents

1. Introduction

Production of vegetable oils has been recognized as a rapidly developing field in plant biotechnology that goes beyond food-based applications. Many kinds of vegetable oils are used in soaps and cosmetics or converted to oleo chemicals that are extensively used to replace petrochemicals in paints, plastics, fuels and lubricants. The demand for biodegradable chemicals applied to industrial products has been increasing and therefore a boost in the production of vegetable oils and fats is needed. Biotechnological approaches including traditional plant breeding and direct genome modification through genetic engineering are crucial tools to increase seed oil production without extending the area of crop cultivation, which has a direct impact on deforestation and competition with food production. Moreover, even diminutive augments in seed oil content reflects in considerable profitability. Despite the unprecedented advances derived from molecular genetics and genomics research on the biochemical pathways of plant lipid metabolism over the last decade, the mechanisms regulating seed oil content are not fully understood. Many aspects of key enzymes are not yet determined even in model plants such as *Arabidopsis thaliana* (Hildebrand et al., 2008). For example, recent studies focusing on intracellular trafficking indicated that compartmentalization of enzyme activities

within the ER membrane represents an additional mechanism adopted by plant cells to control oil production and may be essential for channeling of particular fatty acids into storage lipids (Dyer and Mullen, 2008).

Nevertheless, manipulation of genes involved in storage lipid biosynthesis has been used to increase accumulation of seed triacylglycerol (TAG), the main component of vegetable oils (Weselake, 2002). It was recently demonstrated that overexpression of plant and fungi genes encoding acyl-CoA:diacylglycerol acyltranferase (DGAT, EC 2.3.1.20), that catalyzes the final assembly of TAG, resulted in small but significant increases of seed oil content in canola and soybean tested under field conditions (Lardizabal et al., 2008; Weselake et al., 2008). Indeed, the level of DGAT activity in developing seeds seems to have a direct effect on the accumulation of TAG (Perry and Harwood, 1993; Cahoon, et al., 2007). Surprisingly, little is known about the molecular mechanisms governing DGAT activity. The most basic information about structure and function of this enzyme is essential for rational designs to increase its performance in oilseeds and have a direct reflection in seed oil content. In view of the biotechnological importance of DGATs from plants and fungi, we summarize some of the structural and functional aspects of these enzymes with particular attention to membrane topology, functional polypeptide motifs and subcellular

localization. We use *in silico* approaches to compare the findings obtained with related enzymes in animal and prokaryotes.

2. Discovery of DGAT

The first proceedings reporting DGAT activity date from the 1950s (Weiss and Kennedy, 1956; Weiss et al., 1960), but the genes encoding DGATs were not isolated until the late 1990s. The first *DGAT* cDNA was cloned by taking advantage of homology between an expressed sequence tag (EST) and an acyl-CoA:cholesterol acyltranferase (ACAT, EC 2.3.1.26), a related enzyme previously isolated by a complementation assay of mammalian cells devoid of cholesterol ester biosynthesis (Chang et al., 1993). The mouse (*Mus musculus*) *DGAT* gene isolated in 1998 encodes a protein, here referred to as MmDGAT1 that is 20% identical to mouse ACAT with the most conserved regions on the C-terminus portion of the enzyme (Cases et al., 1998). A plant *DGAT* gene was consequently isolated through the characterization of the locus *TAG1* in an *A. thaliana* EMS-induced mutant (AS11) with altered seed fatty acid composition and decreased DGAT activity (Katavic et al., 1995). The locus *TAG1* contains a 3.4kb gene encoding a polypeptide showing 41% identity with MmDGAT1 (Zou et al., 1999). The polypeptide encoded by *TAG1* (AtDGAT1) exhibits DGAT activity when expressed in yeast and can complement DGAT function in AS11

(Jako et al., 2001). *DGAT* genes from fungi were identified through protein purification, an approach that was previously not successful with other DGATs, perhaps because of their membrane association. Polypeptides exhibiting DGAT activity were purified from lipid bodies of *Umbelopsis ramanniana,* formerly known as *Mortierella ramanniana* (Lardizabal et al., 2001). These *DGATs* shared little or apparently no homology with the previous *DGAT* genes, and therefore were classified as *DGAT2*. Curiously, genes homologous to *DGAT1* have not been found in fungi genomes, although it has been suggested that yeast *ACATs* (*ARE1* and *ARE2* in *Saccharomyces cerevisiae*) represent *DGAT1* orthologues in these organisms as they also display minor DGAT activity (Yen et al., 2008).

Several lines of evidence suggest that DGAT1 belongs to a class of enzymes with acyl-CoA transferase activity that can utilize different acceptors in addition to diacylglycerols. For example, MmDGAT1 also possesses acyl-CoA:retinol acyltransferase (ARAT, EC 2.6.1.57) activity (Yen et al., 2005), while an *A. thaliana* acyl-CoA:fatty alcohol acyltransferase (wax ester synthase, WSD1) also displays DGAT activity *in vitro* (Li et al., 2008). In the case of DGAT2 a similar scenario is observed. In animals, *DGAT2* belong to a gene family with seven members in humans (Cases et al., 2001). Three of these genes encode polypeptides with acyl-CoA monoacylglycerol acyltransferase (MGAT,

EC 2.3.1.22) activity (Yen et al., 2002; Yen and Farese, 2003; Cheng et al., 2003). Two additional members display acyl-CoA:wax alcohol acyltransferase (AWAT, EC 2.3.1.75) activity, which is analogous to WSD1 (Turkish et al., 2005).

Orthologues of *DGAT1* and *DGAT2* have been identified through DNA homology in many other organisms and are widely distributed in eukaryotes. Currently, a relatively wide collection of *DGAT* genes is available, which facilitates more detailed studies of enzyme structure and function through bioinformatics approaches. Many of these genes have been functionally characterized in recombinant systems as described in **Table 1**.

In prokaryotes a bifunctional WS/DGAT was identified in *Acinetobacter calcoaceticus* (Kalscheuer and Steinbuchel, 2003). WS/DGAT has no sequence similarity to DGAT1, DGAT2 or any of the related acyltransferases from eukaryotes. Another non-homologous DGAT, referred to as AhDGAT, was characterized in peanuts (Saha et al., 2006). Unlike other eukaryote enzymes, AhDGAT was purified from the soluble fraction of developing peanuts. Biosynthesis of TAG in the cytosol has been previously reported in a 10S multi-enzyme complex from the oleaginous yeast *Rhodotorula glutinis* (Gangar et al, 2001). Whether this soluble yeast DGAT and AhDGAT compose a novel class DGATs is yet to be demonstrated.

3. Membrane topological organization of DGATs

The pattern in which a protein transverses the membrane bilayer is essential for elucidating the dynamics of the protein structure. DGAT1 and DGAT2 contain hydrophobic segments which are generally believed to constitute transmembrane domains (**Figure 1**). DGAT1 displays more hydrophobic segments than DGAT2, which indicates a different topology and may relate to different physiological roles in TAG biosynthesis (Yen et al., 2008). Few experimental studies on DGAT topological organization in plants and yeast are available and therefore we will mainly rely on *in silico* approaches to predict transmembrane segments and the orientation in the membrane bilayer.

A variety of web-based tools is available for predicting the topology of membrane proteins. Since only a few membrane proteins from bacteria are known to be beta-barrel shaped so far, the prediction algorithms are mostly developed for alpha-helical membrane proteins. Generally, five types of techniques have been using in these programs: hydrophobicity analysis combined with the positive inside rule (TMpred and SOSUI), multiple sequence alignment (ConPredII and TOPCONS), model-recognition approach (MEMSTAT3, TMHMM and HMMTOP) and support vector machine technique (SVMtm) (Persson, 2006). An evaluation of the reliability of these methods indicated that a consensus prediction

and model-based methods are best performing (Moller et al, 2001; Ikeda et al, 2002). The application of these algorithms for the prediction of TM domains in DGAT1 is described in **Table 2** using AtDGAT1 and MmDGAT1 as models. For AtDGAT1, nine of the ten putative transmembrane domains are highly conserved among almost all the prediction results except the domains at 276-299 and 314-337 of AtDGAT1 and 251-276 and 285-312 of MmDGAT1 (highlighted TM5 and TM6 in **Table 2**). A model of nine membrane-spanning topology agrees with our initial study on DGAT1 from *Brassica napus* (Foroud, 2005). In this work, protease mapping data showed that the region between 276 to 299 in BnDGAT1 (corresponding to TM5) is in the cytosol, in agreement to most of the prediction algorithms described in **Table 2**. Recent studies on DGAT1 from *Vernicia fordi* (tung tree) and *B. napus* indicated that the N-terminus face the cytosolic side (Shockey et al, 2006; Weselake et al., 2006) as predicted by most algorithms in **Table 2**. The interaction of N-terminus with lipid substrates in the cytoplasm may lead to a regulatory role of N-terminal region (Siloto et al, 2008) and there are several lines of evidence not only from *B. napus* DGAT1 but mammalian DGAT1 and ACAT1 that favors this hypothesis (Cheng et al, 2001; Yu et al., 1999; Weselake et al, 2006). According to the work on VfDGAT1 the C-terminus of DGAT1 is also proved to orient toward cytosolic side, indicating an even number

of membrane-spanning regions. This result disagrees with a nine transmembrane topology model and therefore further experimental testing will required to examine the hypothesis of eight transmembrane domains.

Compared to DGAT1, DGAT2 is less hydrophobic, having a lower number of transmembrane domains and therefore a less intricate topology. The membrane topology of MmDGAT2 was experimentally determined revealing two transmembrane domains that are closely associated or a single hydrophobic domain embedded in the membrane bilayer (Stone et al., 2006). The first transmembrane domain (TM1) of MmDGAT2 and ScDGAT2 was ubiquitously predicted but the second (TM2) was identified by only a few algorithms (**Table 3**). Since the homology of DGAT2 from different organisms is lower than that of DGAT1, it is possible that ScDGAT2, which has a distinct hydropathy plot, could have a different topology from other fungi DGAT2s. This could be demonstrated by the prediction results of *Schizosaccharomyces pombe* SpDGAT2 (**Table 3**). Interestingly, the prediction of N-terminus orientation seems to be related to the length of the predicted N terminal tail. DGAT2s with putative long tails are intended to face toward the cytosol which agrees with work on VfDGAT2 and MmDGAT2 (Shockey et al 2006; Stone et al., 2006). The same conclusion, however, cannot be made for DGAT2s with short tails.

4. Alignment of DGAT1 polypeptides

DGAT1 polypeptides are typically characterized by a hydrophilic N-terminus sequence followed by a number of hydrophobic stretches constituting potential transmembrane domains as previously discussed. The total number of predicted transmembrane domains in DGAT1 can vary according to the sequence and the algorithm used as shown above. When the sequences are aligned, however, many of these potential transmembrane domains are found in the same positions in most DGAT1 (**Figure 2**). The first four transmembrane domains on the first half of the sequences and the last three transmembrane domains on the C-terminus are separated by short polar loops. Between these groups are two possible membrane spanning regions that are separated by longer hydrophilic stretches. Here we will consider these nine potential transmembrane domains as landmarks to describe conserved motifs in DGAT1 acknowledging, however, that an experimental approach is required to verify these assumptions. We will also use the sequence of *A. thaliana* DGAT1 to describe the exact position of each motif.

An overview of the DGAT1 alignment from 30 different organisms indicates several conserved residues with about 7% of identical residues among plant and animal sequences. The hydrophilic N-terminus is composed of an

average of 115 and 80 residues in plants and animals, respectively and is the least conserved region in DGAT1. An alignment of the N-terminal portion of DGAT1 from a broad range of organisms revealed a cluster of arginines in the first 30 residues (**Figure 3**). The region comprising twenty positions preceding the first hydrophobic domain is also conserved and contains the motifs PAHRXXXESPLSSDAIFXQ and SLFSXXSGFXN which are conserved in plants and animals, respectively. Other divergences discriminating DGAT1 from plants and animals include a serine at position 131 of AtDGAT1 conserved in plants and absent in animals DGAT1, and the motif WVXRQ in plants, corresponding to FL($^{L}/_{I}$)($^{R}/_{K}$)R in animals. These differences can be also observed in more ancient organisms such as *Toxoplasma gondii* and *Physcomitrella patens*. The long loop between the fourth and fifth transmembrane domains (between positions 260 to 278 of AtDGAT1) shows remarkable variability amongst all DGAT1s. Following this region lies the most conserved uninterrupted sequence of DGAT1 comprising the motifs PTLCYQXSYPR in plants and PTLCYEXXFPR in animals, preceding the fifth predicted transmembrane domain between positions 292 to 297 of AtDGAT1.

5. Alignment of DGAT2 polypeptides

DGAT2 polypeptides, in comparison to DGAT1, display fewer potential transmembrane domains and higher sequence divergence. An alignment of DGAT2 sequences from 20 organisms covering plants and animals previously described in DGAT1, indicate approximately 5% of identical residues. Inclusion of 16 fungi sequences in this group decrease the identity to only 2.3%. A higher divergence might pose difficulties for identification of novel members of DGAT2 through sequence homology.

At least one transmembrane domain can be predicted for every DGAT2, but usually two transmembrane domains are conserved in the N-terminus portion and separated by a small loop (**Figure 4**). This hydrophobic region is definitely very important as its removal results in lack of activity in ScDGAT2 (unpublished). An experimental approach indicated that in MmDGAT2 the only membrane-spanning region is composed of two transmembrane domains separated by a small loop that could be also interpreted as a single hydrophobic region embedded in the membrane (Stone et al., 2006). This region will be used as a landmark and the positions of conserved motifs also will be indicated in the UrDGAT2a polypeptide from *U. ramanniana*. The N-terminus portion preceding the transmembrane domains is quite variable length and is usually smaller in

animals and plants (with 38 and 30 residues in average) when compared to fungi (with 100 residues in average). The most conserved region in DGAT2 encompasses the motif RXGFX($^{K}/_{R}$)XAXXXGXX($^{L}/_{V}$)VPXXXFG($^{E}/_{Q}$) located approximately 150 residues after the second transmembrane domain (positions 259 to 281 of UrDGAT2a). Other conserved residues are the motif GGXXE (positions 204 to 208 in UrDGAT2a) and a phenylalanine, an arginine and a proline in positions 164, 170 and 293 of UrDGAT2a, respectively. In addition, the motif YXXXXXHPHG is conserved in sequences from animals and fungi (positions 121 to 129 of UrDGAT2a) corresponding to YXXXXXEPH$^{S}/_{G}$ in plants. Preceding this motif is situated one of the most striking divergences in DGAT2 alignment, a hydrophilic segment of approximately 41 residues present in sequences from some fungi but absent in plants and animals. This region, corresponding to positions 144 to 185 of ScDGAT2, is also found as a much larger segment (158 residues) in *Yarrowia lipolytica* DGAT2. This hydrophilic segment, although non-essential, was demonstrated to modulate the enzyme activity of ScDGAT2 (unpublished results). Because this segment precedes a highly conserved motif, it is possible that it might represent a specialized function in DGAT2 from certain fungi.

6. Structure of DGAT genes

The architecture of genes encoding DGAT is largely available from whole genome sequence databases or, as in the case of *V. fordii*, from sequencing of the respective genomic regions. In mammals, genes encoding DGAT1 share a similar architecture of 17 exons mostly grouped in the 3' portion. A *DGAT1* representative from invertebrates (*Caenorhabditis elegans*), however, shows an unrelated distribution with only 7 exons (**Figure 5A**). In plants, DGAT1 genes from *A. thaliana*, *Medicago truncatula*, *Zea mays* and *V. fordii* are composed by 16 exons, while DGAT1 from *Oryza sativa* contain 14 exons (**Figure 5B**). The first exon of plant *DGAT1* genes comprises the largest coding sequence and encodes the hydrophilic N-terminus. Curiously, the last codon from the first exon of these genes encodes a glutamine in the same position of the alignment (motif IFX**Q**), denoting the end of the hydrophilic N-terminus and start of the first predicted membrane spanning region (**Figure 3**). The hydrophilic N-terminus is the most variable sequence of DGAT1 polypeptides and, therefore, it is possible that segregation of this sequence in the first exon might have been used as an evolutionary mechanism to delimit variability in this region of the gene. This pattern was not observed in *DGAT1* sequences from animals. *DGAT2* genes show a dissimilar structure to *DGAT1*. Mammalian *DGAT2* genes share a common

architecture with 8 exons while the gene from *C. elegans* has only two exons (**Figure 5C**). In plants *DGAT2* genes have 8 exons in *A. thaliana* and 10 exons in *V. fordii* and *O. sativa* (**Figure 5D**).

7. Functional motifs in DGAT1

Most of the information available on the structure and function of DGATs is derived from comparisons of homologous enzymes. Alignments of polypeptide sequences encoding acyl-CoA-dependent acyltransferases from diverse organisms indicated a conserved histidine and an aspartic acid in the configuration **H**XXXX**D.** Substitution of the conserved histidine in the bifunctional enzyme 2-acyl-glycerophosphoethanolamine acyltransferase / acyl-acyl carrier protein synthase (Aas, EC 2.3.1.40 and 6.2.1.20, respectively) resulted in lack of acyltransferase activity (Heath and Rock, 1998). Substitution of the aspartic acid residue also resulted in significantly less activity. It was suggested that the histidine operates as a general base to abstract the proton from the hydroxyl group of *sn*-1 glycerol-3-phosphate facilitating nucleophilic attack on the thioester bond of acyl-CoA. The aspartic acid would work in a charge relay system to increase the nucleophilicity of the hydroxyl group. This mechanism could be used by other acyltransferases, including DGAT. In fact a similar motif (**HH**XXX**DG**) is conserved in DGATs from prokaryotes (Daniel et al., 2004). In eukaryotic

DGAT1, the motif **H**XXX**D** can be found closely after the fourth predicted transmembrane domain in DGAT1 from plants (positions 257 to 261 of AtDGAT1). Similarly, the motif **H**XXXX**D** is found in a region preceding the fifth predicted transmembrane domain of a few plants such as *A. thaliana*, *B. napus*, *Ricinus communis* and *V. fordii* (positions 342 to 347 of AtDGAT1) (**Figure 6**). These motifs, however, are not conserved in animals and therefore might not compose the catalytic site of DGATs. Jako et al. (2001) identified the consensus sequence N(S/A/G)**R**(L/V)(I/F/A)(I/L)**E**N(L/V) in AtDGAT1 and proposed that the invariant arginine and glutamic acid on positions 149 and 153, could have analogous function to the histidine and aspartic acid residues, respectively. This region is highly conserved in all organisms including more ancient eukaryotes (*T. gondii* and *P. patens*) (**Figure 6**). These residues are present in the interface between a putative transmembrane domain and the adjacent hydrophilic loop, which would create an amphipathic environment for the substrates of DGAT. Moreover, DGAT1 is recognized as a member of a large protein family of membrane-bound O-acyltransferases known as MBOAT (NCBI domain ID pfam03062; Hofmann, 2000). Other members of the MBOAT family catalyze O-acylation reactions transferring acyl chains onto hydroxyl or thiol groups of lipids and proteins. For example, ACAT transfers an acyl chain from

acyl-CoA to cholesterol, forming cholesteryl esters (Chang et al., 1993) while skinny hedgehog (ski) protein transfers a palmitoyl group onto cysteine residues of other proteins (Chamoun et al., 2001). The MBOAT family is characterized by a hydrophobic region (positions 234 to 509 of AtDGAT1) which contains a conserved asparagine (position 410 in AtDGAT1) and histidine (position 447 in AtDGAT1) (**Figure 6**). It has been proposed that these residues could be involved in the catalytic activity. For example, this conserved histidine has been demonstrated to be a key residue for human ACAT1 activity (Guo et al., 2005). Whether any of these regions contribute to the catalytic site of DGAT1 is yet to be experimentally tested. Interestingly, *sn*-1 glycerol 3-phosphate acyltransferase (GPAT, EC 2.3.1.15) and lysophosphatic acid acyl transferase (LPAAT, EC 2.3.1.51), which are also membrane-bound O-acyltransferases catalyzing the first two acylation steps of TAG biosynthesis are not classified as MBOAT members, suggesting that these enzymes might not share similar catalytic sites. It is also possible that these residues could act as supplementary catalytic sites being involved in other enzyme activities besides DGAT, such as ARAT and ACAT.

Other putative active sites in DGAT1 include the substrate binding sites. Sequences of DGAT1 from several plants indicate the presence of a putative diacylglycerol/phorbol ester binding motif that is apparently absent in ACATs

(Zou et al., 1999; Nykiforuk et al., 2002; Xu et al., 2008). Phorbol esters such as phorbol-12-myristate-13-acetate (PMA) are commonly known to mimic diacylglycerols. The putative diacylglycerol/phorbol ester-binding motif present in the positions 414 and 424 of AtDGAT1 forms the consensus **H**XXXX**RH**XXX**P** in DGAT1 from plants and animals. Xu et al. (2008) demonstrated that substitution of a phenylalanine to an arginine in position 439 of TmDGAT1 that is 16 positions after the predicted motif resulted in loss of DGAT activity. This could be a result of alterations in DAG interaction with DGAT. But, because this phenylalanine is positioned at a predicted transmembrane domain, substitution to a charged residue could also have structural implications. Acyl-CoA has been shown to interact with a recombinant N-terminal segment of BnDGAT1 and MmDGAT1 (Weselake et al., 2000; Weselake et al., 2006; Siloto et al., 2008). The N-terminus sequence is highly variable, except for a region of 20 residues preceding the first hydrophobic domain which shows remarkable conservation among plants and animals. Many of these variations, however, represent amino acid residues with similar properties, which could explain the acyl-CoA binding properties of DGAT1 from *B. napus* and *M. musculus*. Several lines of evidence suggest that acyl-CoA interaction with the hydrophilic N-terminus of DGAT1 regulates this enzyme allosterically. First, there is positive

cooperativity exhibited for binding of 22:1-CoA in mouse and canola DGAT1 (Weselake et al., 2000; Siloto et al., 2008). Second, enzymes that are allosterically regulated often form multimeric complexes to achieve cooperativity and the N-terminus of DGAT1 assists on the formation of dimers and tetramers as demonstrated for BnDGAT1and HsDGAT1, respectively (Weselake et al., 2006; Cheng et al., 2001). For example, ACAT1 self-associates through the N-terminus which also plays a regulatory role in this enzyme (Guo et al., 2001; Yu et al., 2002). Third, the acyl-CoA binding motif is not essential for enzyme activity as the removal of the N-terminus of RcDGAT1 results in a polypeptide with substantial enzyme activity, indicating that this is not the exclusive region to interact with acyl-CoA (unpublished data). Indeed, the fourth conserved block in GPATs and LPAATs, as described by Lewin et al. (1999), contains an invariant proline that has been proposed to participate in acyl-CoA binding. This proline was identified in plant DGAT1 polypeptides on the third predicted transmembrane domain corresponding to position 224 of AtDGAT1 and is in fact conserved in DGAT1 from all organisms. Substitution of this proline to an arginine in TmDGAT1 abolished DGAT activity corroborating to the idea that this residue has a functional role (Xu et al., 2008). Another possible acyl-CoA binding site was proposed to be closely associated with the motif FYXDWWN in

ACATs (Yen et al., 2008). This motif is present in DGAT1 and shows remarkable conservation with exception to CeDGAT1, where the second tryptophan is substituted by a phenylalanine. This motif locates on the loop preceding the third last putative transmembrane domain, relatively distant from the proline residue previously discussed, but near the asparagine residue conserved in MBOAT members. The paired tryptophans in this motif are a rare combination and have been previously demonstrated to participate in cholesterol binding. Guo et al. (2001) demonstrated, however, that substitution of the conserved tyrosine to alanine in yeast ACAT1 resulted in decreased affinity for acyl-CoA. Substitution of this same residue in TmDGAT1 (Y392A) resulted in decreased enzyme activity while a double mutation in tyrosine and tryptophan (Y392G / W395G) completely abolished enzyme activity (Xu et al., 2008).

Other putative functional domains predicted in DGAT1 include a leucine zipper and phosphorylation sites although is not yet clear whether these regions are important in the function, structure or regulation of DGAT1. A putative leucine zipper motif was described in several DGAT1 from plants (Bouvier-Nave et al., 2000a; Nykiforuk et al., 2002). For example, in AtDGAT1 five leucines (L222, L229, L236, L243, L250), are consecutively spaced by six residues forming a classic leucine zipper (Hobbs et al., 1999). This leucine zipper, that

might mediate interactions with other proteins, is present in a number of DGAT1 from plants but not from animals. Several studies indicated the presence of multiple potential phosphorylation sites in DGAT1 (Hobbs et al., 1999; Nykiforuk et al., 2002; He et al., 2004). Some of these sites are conserved in plant DGATs, such as the protein kinase C sites in the loop between the first and second transmembrane domain (positions 169 to 171 and 172 to 175 of AtDGAT1) and the caseine kinase II sites (positions 254 to 257 and 403 to 406 of AtDGAT1). In addition a tyrosine kinase site (positions 386 to 393 of AtDGAT1) is conserved in DGAT1 from plants and animals. This site overlaps with the FYXDWWN motif previously discussed as a putative acyl-CoA binding site. Although substitution of the conserved tyrosine to alanine in yeast ACAT homologue resulted in lower affinity to acyl-CoA, phospholylation could not be directly detected (Guo et al., 2001). Regulation of DGAT1 activity through phosphorylation is compelling not only because this is a common mechanism to control enzyme activity in eukaryotes but also because DGAT can scavenge DAG, an important molecule involved in phosphorylation signaling cascades (Carrasco and Merida, 2007). For example, the affinity of DAG to C1 domains of DAG kinase is modified by phosphorylation of residues situated close to this motif (Thuille et al., 2005). In addition, the fact that DGAT is expressed in vegetative tissues suggests that it can

have additional roles beyond oil biosynthesis in seeds (Lu et al., 2003). Substitution of serine at position 168 in RcDGAT1 that corresponds to a protein kinase C site previously described resulted in a significant decrease in the enzyme activity (unpublished).

8. Functional motifs in DGAT2

The motifs previously described for DGAT1 cannot be found in DGAT2 sequences likely due to the little homology between the DGAT1 and DGAT2. In fact, little is known about functional motifs of DGAT2. Stone et al. (2006) identified the conserved motif HPHG in positions 161 to 164 of MmDGAT2 as an important region for DGAT activity. Substitution of these residues, forming the sequences **A**PHG, H**G**HG, HP**A**G and **AGA**G resulted in significant reduction of enzyme activity. More specifically, the histidine at position 163 of MmDGAT2 appeared to play a more important role for the enzyme function, which agrees with our mutagenesis work on ScDGAT2 (unpublished). This region is conserved in animal and fungi DGAT2, but in plants this motif is found as EPH$^{S}/_{G}$. Substitution of the glutamic acid to a histidine residue in plant DGAT2 did not result in appreciable effect, but replacement of the motif HPHG in ScDGAT2 with residues EPHS found in plant DGAT2 resulted in loss of enzyme activity (unpublished). This indicates an important divergence on the structure/function of

DGAT2 from fungi and plants. In addition, the motif FLXLXXX^{n} (n indicates a nonpolar residue) was indicated as a putative neutral lipid binding domain in MmDGAT2 (Stone et al., 2006). Substitution of phenylalanine (position 80) and leucine (position 81) residues to alanine residues resulted in decreased DGAT activity. Substitution of the leucine in position 83 to an alanine resulted in lack of activity. This motif, present in the first predicted transmembrane domain of MmDGAT2 (positions 80 to 87) is conserved in vertebrate DGAT2 but not in plants or fungi orthologues. Substitution of the corresponding phenylalanine and leucine (positions 71 and 73) in ScDGAT2 results in ~50% of the wild-type activity (unpublished). This same motif contains the putative membrane lipid attachment LGVAC found in prokaryotes through the interaction between the sulfhydryl group of a cysteine residue (position 87 of MmDGAT2) and DAG. Substitution of this cysteine to a serine in MmDGAT2 did not reduce DGAT activity, indicating that it does not function as a lipid attachment site. In fact, substitution of all cysteine residues in ScDGAT2 to alanine residues did not disrupt DGAT activity indicating that this mechanism of DAG interaction is not present or at least essential in this enzyme (unpublished data). In addition substitutions on the conserved motif YFP located close to the transmembrane

domains (positions 104 to 106 of UrDGAT2A) resulted in significant decreases in activity in ScDGAT2.

9. Subcelluar localization of DGATs

To better elucidate the role of DGATs in cellular processes, their spatial location has been studied in different plants. In numerous earlier studies DGAT location has been a subject of discrepancy whether it is associated with endoplasmic reticulum (ER) or oil bodies (Lung and Weselake, 2006). This debate could be the result of technique limitations as the general approach used in these studies was subcelluar fractionation combined with enzyme assay in which cross-contamination can occur. For instance, in the study of germinating soybean cotyledon, the purified oilbodies also exhibited activities for ER markers (Settlage *et al.*, 1995). This could be explained by association between oilbodies and ER (Cao and Huang, 1986; Settlage *et al.*, 1995). Lacey *et al* (1996) applied different organelle markers to rule out the possible contamination in the assay and clearly demonstrated that *B. napus* DGAT is associated with ER. Similarly, Cao and Huang (1986) were able to localize maize DGAT in the rough ER (RER) by taking advantages of protein markers as well as the attachment of RER with polysomes in the presence of Mg^{2+} during fractionation. Actually the ER is regarded to be the main site for TAG synthesis and microsomal fractions from

developing seeds of many plants as well as plant cultured cells have been extensively utilized for enzyme assays (Browse and Somerville, 1991; Weselake 2005). Using more dependable techniques such as green fluorescent protein (GFP)-tagging and immunofluorescence, Shockey *et al.* (2006) have demonstrated that tung tree DGAT1 and DGAT2 are localized in the ER. Localization of both DGATs is dependent on a C-terminal ER retrieval motif. In VfDGAT1 the ER retrieval sequence YYHDL is part of the motif LLYYHDXMN conserved in all plant DGAT1. The ER retrieval domain in VfDGAT2 comprises the sequence LKLEI, where the two leucines are conserved in other DGAT2 sequences. Removal of the corresponding regions through C-terminus truncations in RcDGAT1 and ScDGAT2 resulted in decreased activity and decreased protein stability, respectively, indicating the importance of the C-terminus portion for both DGATs (unpublished). Interestingly, VfDGAT1 and VfDGAT2 do not co-localize in the ER and therefore it is plausible that these polypeptides have distinct interactions with other proteins in the ER membrane. Mounting evidence based on studies with animals and plants indicate that DGAT1 and DGAT2, although catalyzing the same enzyme activity, have distinct physiological functions (Yen et al., 2008; Shokey 2006; Burgal 2008). In addition to the ER, DGAT activity was also found in chloroplasts of spinach leaves (Martin and

Wilson, 1984) and more recently, Kaup et al. (2002), identified DGAT1 in the chloroplasts of senescing *Arabidopsis* leaves through immunoblotting. The mechanisms by which AtDGAT1 is transported to the chloroplast are yet to be determined.

In yeast, biochemical studies with *S. cerevisiae* indicated that DGAT activity is mainly in lipid droplets (Sorger and Daum, 2002). Indeed, DGAT2 in *U. rammaniana* was purified from the lipid particle fractions (Lardizabal et al., 2001). In addition, two subcellular localization datasets generated by proteomic studies of *S. cerevisiae* indicated that ScDGAT2 localizes in ER and lipid droplets (Huh et al., 2003; Natter et al., 2005). In addition, recombinant expression of ScDGAT2 in a yeast strain devoid of TAG biosynthesis, indicated that ScDGAT2 localizes in the microsomal fraction as an integral membrane protein (unpublished). Due to the presence of conserved transmembrane domains, it is expected that yeast DGAT2 localizes in the ER. The mechanisms involved in its transfer from the ER to lipid droplets, however, were not yet determined. *S. cerevisiae* lipid droplets contain a small fraction of proteins (Leber et al., 1994) when compared to oil bodies, the structurally related organelles in plants that are coated by oleosins (Tzen et al., 1993). The mechanism of oilbody targeting in oleosins has been well studied and is assisted by a long hydrophobic domain

composed of two chains that are separated by a motif with three conserved prolines (proline knot motif) that supposedly folds the domain, resulting in an unusual topological structure where the hydrophilic N- and C-termini face the cytoplasm (Tzen et al., 1992; Abell et al., 2004). Analysis of yeast DGAT2 revealed that the two potential transmembrane domains are separated by a very small loop. This region contains three prolines that are conserved in most sequences (**Figure 7**). Such similarities could explain the transfer of DGAT2 to lipid droplets although such hypothesis needs to be experimentally verified. A study on murine DGAT2 has shown that the enzyme localizes on the ER and it transfers to near the surface of lipid droplets when oleic acid is provided to drive TAG biosynthesis (Stone et al., 2009). Interestingly, determination of murine DGAT2 topological structure also showed both termini facing the cytosol and the possible presence of two adjacent transmembrane domains (Stone et al., 2006). DGAT2 from other organisms have a similar structure. VfDGAT2 has both N- and C-termini facing the cytosol with two predicted transmembrane domains separated by a small loop with a conserved proline (Shockey et al., 2006).

10. Conclusions and future research

Considerable progress has been achieved towards our understanding of DGATs and their involvement in the biosynthesis of TAG over the last decade.

Many important aspects of the molecular mechanisms coordinating the catalytic activity, however, remain unclear. Most interestingly, DGAT1 and DGAT2 are unrelated polypeptides and yet catalyze the same reaction. Are the mechanisms involved in the acyltransferase catalytic function similar in DGAT1 and DGAT2? Are there any relationships between these enzymes that have not been identified with the current alignment algorithms? In an evolutionary perspective, did these enzymes evolve separately to catalyze the same reaction or do they have a common ancestor? These are some unanswered questions that require more fundamental research on DGATs. It would be valuable to have insights into the three dimensional structure of DGATs as it would help to resolve some of these doubts.

Most of the information on putative structure-function relationships in DGATs has been deduced using bioinformatics approaches. The conclusions obtained with such approaches are valuable but still require experimental validation. Considerable progress has been made in shedding light on the topological organization of murine DGAT2 (Stone et al., 2006). It would be interesting to conduct similar experiments with a fungal DGAT2, particularly with ScDGAT2. This polypeptide contains unique characteristics as previously discussed making it an interesting candidate for structural studies. Moreover, the

correct topology of DGAT1 should be experimentally evaluated to determine whether the nine or ten transmembrane model is the correct one. In addition, to enhance our knowledge on DGAT catalytic activity, broader studies involving site-directed mutagenesis should be performed to identify functional regions. Currently, two studies have been conducted with a plant DGAT1 (*Tropaeolum majus*) and an animal DGAT2 (*M. musculus*) evaluating the influence of only a few residues (Xu et al., 2008; Stone et al., 2006). The polypeptide alignments presented in this chapter indicate the presence of multiple sites that could be involved on the catalytic activity of DGAT. This type of research could greatly advance if random mutagenesis techniques such as directed evolution or site-saturation mutagenesis could be applied to DGATs. One of the obstacles associated with such large-scale experiments is that standard methods to accurately measure DGAT activity require a laborious assay with radio-labeled substrates (Coleman et al., 1992). Due to association of DGATs with membranes, enzyme assays typically use microsomal fractions obtained through ultracentrifugation which greatly decreases the throughput of the assay. Recently we have demonstrated two assays to detect and measure DGAT activity in high throughput scale (Siloto et al., 2009). Further development of such assays would definitively enhance our knowledge about the molecular mechanisms involved in

DGAT activity. Furthermore, this is an attractive field for plant biotechnology for improving the performance of DGATs from plants and fungi which have been already used to increase oil content in seeds (Weselake et al., 2008; Lardizabal et al., 2008).

Acknowledgements: The authors are grateful to the following grant agencies for the funding they provided to us in support of our DGAT characterization work: AVAC Ltd, Alberta Advanced Education and Technology, the Alberta Agricultural Research Institute, the Alberta Crop Industry Development Fund, the Canada Research Chairs Program, Genome Alberta and the Natural Sciences and Engineering Research Council of Canada.

Abbreviations: EMS, ethyl methanesulfonate; TAG, triacylglycerol; DGAT, acyl CoA:diacylglycerol acyltransferase; EST, expressed sequence tag; acyl CoA:cholesterol acyltransferase, ACAT; ARAT, acyl-CoA:retinol acyltransferase; MGAT, acyl-CoA monoacylglycerol acyltransferase; TM, transmembrane; MBOAT, membrane-bound O-acyltransferases; PMA, phorbol-12-myristate-13-acetate; ER, endoplasmic reticulum; RER, rough ER; GFP, green fluorescent protein.

The two-character code preceding each *DGAT* gene and polypeptide indicate the organism of origin as follows:

Aa, *Aedes aegypti*; Ac, *Aspergillus clavatus*; Ah, *Arachis hypogaea*; An, *Aspergillus niger*; Ao, *Aspergillus oryzae*; At, *Arabisopdid thaliana*; Bf, *Branchiostoma floridae*; Bn, *Brassica napus* ; Bt, *Bos Taurus*; Ce, *Caenorhabditis elegans*; Ci, *Coccidioides immitis*; Cn, *Cryptococcus neoformans*; Cr, *Chlamydomonas reinhardti*; Dm, *Drosophila melanogaster* ; Dd, *Dictyostelium discoideum*; Dr, *Danio rerio*; Ea, *Euonymus alatus*; Gm, *Glycine max*; Gz, *Gibberella zeae*; Hs, *Homo sapiens*; Jc, *Jatropha curcas*; Lb, *Laccaria bicolor*; Md, *Monodelphis domestica*; Mm, *Mus musculus*; Mt, *Medicago truncatula*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*; Nf, *Neosartorya ischeri*; Nt, *Nicotiana tabacum*; Nv, *Nematostella vectensis*; Oe, *Olea europaea*;

Os, *Oryza sativa*; Ot, *Ostreococcus tauri*; Pf, *Perilla frutescens*; Pm, *Penicillium marneffei*; Pn, *Phaeosphaeria nodorum*; Pp, *Physcomitrella patens*; Pt, *Populus trichocarpa*; Ps, *Picea sitchensi*; Rc, *Ricinus communis*; Rg, *Rhodotorula glutinis*; Rn, *Rattus norvegicus*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ss, *Sus scrofa*; Ta, *Trichoplax adhaerens*; Tg, *Toxoplasma gondii*; Tm, *Tropaeolum majus*; Ts, *Talaromyces stipitatus*; Um, *Ustilago maydis*; Ur, *Umbelopsis ramanniana*; Vf, *Vernicia fordii*; Vg, *Vernonia galamensis*; Vv, *Vitis vinifera*; YL, *Yarrowia lipolytica*; Zm, *Zea mays*.

Table 1: Eukaryotic *DGAT*s functionally tested in recombinant organisms.

cDNA	Original organism	Host used for expression and relevant genetic markers	Reference
HsDGAT1	*H. sapiens*	*S. cerevisiae* 12501 (*dga1⁻Δ*)	Inokoshi et al., 2009
AtDGAT1	*A. thaliana*	*B. napus*	Weselake et al., 2008
BnDGAT1	*B. napus*	*B. napus*	Weselake et al., 2008
TmDGAT1	*T. majus*	*S. cerevisiae* H1246 (*are1⁻ Δ, are2⁻ Δ, dga1⁻ Δ, lro1⁻ Δ*), *A. thaliana* and *B. napus*	Xu et al., 2008
ZmDGAT1	*Z. mays*	*Z. mays* and *S. cerevisiae* (*dga1⁻ Δ, lro1⁻ Δ*)	Zheng et al., 2008
VgDGAT1	*V. galamensis*	*S. cerevisiae*	Yu et al., 2008
VfDGAT1	*V. fordii*	*S. cerevisiae* SCY1998 (*dga1⁻ Δ, lro1⁻ Δ*)	Shockey et al., 2006
AhDGAT	*A. hypogaea*	*E. coli*	Saha et al., 2006
MmDGAT1	*M. musculus*	COS7-cells (*C. sabaeus*)	Yen et al., 2005
EaDGAT1	*E. alatus*	*S. cerevisiae* H1266 (*are2⁻ Δ, dga1⁻ Δ, lro1⁻ Δ*)	Milcamps et al., 2005
AtDGAT1	*A. thaliana*	*S. cerevisiae* H1266 (*are2⁻ Δ, dga1⁻ Δ, lro1⁻ Δ*)	Milcamps et al., 2005
RcDGAT1	*R. communis*	*S. cerevisiae*	He et al., 2004
TgDGAT1	*T. gondii*	*S. cerevisiae* SCY910 (*are1⁻ Δ, are2⁻ Δ*)	Quittnat et al., 2004
HsDGAT1	*H. sapiens*	McA-RH7777 cells (*R. norvegicus*)	Liang et al., 2004
BnDGAT1	*B. napus*	*P. pastoris*	Nykiforuk et al., 2002
MmDGAT1	*M. musculus*	Sf9 insect cells (*S. frugiperda*)	Cases et al., 2001
AtDGAT1	*A. thaliana*	*A. thaliana* AS11	Jako et al., 2001
AtDGAT1	*A. thaliana*	*S. cerevisiae* SCY 062	Bouvier-Nave et al., 2000a
CeDGAT1	*C. elegans*	*S. cerevisiae* SCY 062	Bouvier-Nave et al., 2000a
NtDGAT1	*N. tabacum*	*S. cerevisiae* SCY 062	Bouvier-Nave et al., 2000a
AtDGAT1	*A. thaliana*	*S. cerevisiae* SCY059 (*are1⁻ Δ, are2⁻ Δ*), *N. tabacum*	Bouvier-Nave et al., 2000b
AtDGAT1	*A. thaliana*	Sf21 insect cells (*S. frugiperda*)	Hobbs et al., 1999
AtDGAT1	*A. thaliana*	*S. cerevisiae* YMN5 (*Slc1⁻ Δ*)	Zou et al., 1999
MmDGAT1	*M. musculus*	H5 insect cells (*T. ni*)	Cases et al., 1998
HsDGAT2	*H. sapiens*	*S. cerevisiae* 12501 (*dga1⁻ Δ*)	Inokoshi et al., 2009
RcDGAT2	*R. communis*	*A. thaliana* and *S. cerevisiae* (*dga1⁻ Δ*)	Burgal et al., 2008
UeDGAT2	*U. ramanniana*	*G. max*	Lardizabal et al., 2008
RcDGAT2	*R. communis*	*S. cerevisiae*	Kroon et al., 2006
MmDGAT2	*M. musculus*	COS7-cells (*C. sabaeus*)	Stone et al., 2006
VfDGAT2	*V. fordii*	*S. cerevisiae* SCY1998 (*dga1⁻ Δ,lro1⁻ Δ*)	Shockey et al., 2006
HsDGAT2	*H. sapiens*	*S. cerevisiae* ScY2051 (*are2⁻ Δ, dga1⁻ Δ, lro1⁻ Δ*)	Turkish et al., 2005
MmDGAT2	*M. musculus*	Sf9 insect cells (*S. frugiperda*)	Cases et al., 2001
HsDGAT2	*H. sapiens*	Sf9 insect cells (*S. frugiperda*)	Cases et al., 2001
CeDGAT2	*C. elegans*	Sf9 insect cells (*S. frugiperda*)	Lardizabal et al., 2001
ScDGAT2	*S. cerevisiae*	Sf9 insect cells (*S. frugiperda*)	Lardizabal et al., 2001
UeDGAT2	*U. ramanniana*	Sf9 insect cells (*S. frugiperda*)	Lardizabal et al., 2001

Table 2. Prediction results for transmembrane domains in AtDGAT1. The polypeptides corresponding to AtDGAT1 and MmDGAT1 were submitted to a number of transmembrane prediction algorithms. The numbers correspond to the presence of each transmembrane (TM) domain.

		ConPredII	TOPCONS	MEMSAT3	HMMTOP	TMHMM	SVMtm	SOSUI	TMpred
AtDGAT1	**TM1**	132-152	132-152	134-152	133-152	133-152	135-149	131-153	131-152
	TM2	176-196	173-193	177-197	177-195	176-195	179-193	175-197	176-193
	TM3	207-227	209-229	206-230	208-227	207-229	208-222	205-227	208-229
	TM4	234-254	234-254	233-252	236-255	234-256	236-251	235-257	234-256
	TM5	277-297	—	—	276-294	—	—	—	280-299
	TM6	316-336	314-334	312-331	319-336	315-337	316-330	—	314-333
	TM7	365-385	362-382	353-371	363-382	363-385	363-382	370-392	366-385
	TM8	433-453	433-453	431-450	434-453	433-455	433-448	432-454	433-451
	TM9	458-478	458-478	453-475	460-479	460-479	457-473	459-481	460-476
	TM10	488-508	488-508	487-510	490-509	491-513	489-503	484-506	487-509
	Orientation	OUT	IN	IN	IN	IN	IN	N/A	OUT
MmDGAT1	**TM1**	95-115	95-115	97-115	96-114	—	97-111	93-115	96-114
	TM2	140-160	137-157	142-161	141-159	137-159	144-158	141-160	137-157
	TM3	173-193	174-194	171-195	172-195	172-194	175-195	171-193	174-198
	TM4	200-220	198-218	200-224	202-220	198-220	200-214	199-221	200-218
	TM5	256-276	—	—	251-269	—	—	—	—
	TM6	—	292-312	285-308	296-312	293-312	—		293-311
	TM7	339-359	334-354	338-360	343-361	342-364	341-357	337-359	343-364
	TM8	—	412-432	410-429	—	412-434	417-431	—	—
	TM9	439-459	434-454	432-456	439-456	439-456	439-454	—	436-456
	TM10	465-485	465-485	464-483	467-484	463-485	467-481	—	463-483
	Orientation	IN	IN	IN	IN	OUT	N/A	OUT	IN

*Orientation of N terminus. Cytosol-"IN". Lumen-"Out". TMs highlighted in gray are not universally predicted. The websites used for each algorithm are: ConPred II, http://bioinfo.si.hirosaki-u.ac.jp/~ConPred2; TOPCONS, http://topcons.net; MEMSAT3, http://bioinf.cs.ucl.ac.uk/psipred/psiform.html; HMMTOP, http://www.enzim.hu/hmmtop; TMHMM, http://www.cbs.dtu.dk/services/TMHMM; SVMtm, http://ccb.imb.uq.edu.au/svmtm; SOSUI, http://bp.nuap.nagoya-u.ac.jp/sosui; TMpred, http://www.ch.embnet.org/software/TMPRED_form.html.

Table 3. Prediction results for transmembrane domains in DGAT2. The polypeptide corresponding to ScDGAT2, MmDGAT2, AtDGAT2 and SpDGAT2 were submitted to a number of transmembrane prediction algorithms. The numbers correspond to the presence of each transmembrane (TM) domain.

		ConPredII	TOPCONS	MEMSAT3	HMMTOP	TMHMM	SVMtm	SOSUI	TMpred
ScDGA2	**TM1**	73-93	62-82	59-77	68-92	70-92	66-97	75-97	72-92
	TM2	—	84-104	80-104	—	—	—	—	—
	TM3	189-209	188-208	—	—	—	193-207	188-210	196-214
	TM4	215-235	—	—	200-224	—	217-231	213-234	216-236
	TM5	293-313	—	—	294-310	—	—	—	—
	TM6	—	—	—	341-359	—	—	—	—
	Orientation	IN	IN	IN	IN	IN	IN	N/A	IN
MmDGAT2	**TM1**	70-90	71-91	61-85	68-92	73-95	73-93	66-88	76-96
	TM2	—	93-113	88-112	—	—	95-109	93-115	—
	TM3	—	—	160-178	—	—	—	—	—
	TM4	—	—	—	—	—	—	—	224-248
	Orientation	IN	IN	IN	IN	IN	OUT	N/A	IN
AtDGAT2	**TM1**	19-39	20-40	18-36	19-42	15-37	19-43	16-38	27-50
	TM2	44-64	43-63	39-60	47-64	39-61	45-60	44-65	—
	TM3	109-129	—	103-121	112-129	—	112-126	—	112-130
	TM4	134-154	—	—	—	—	138-153	136-153	134-154
	Orientation	OUT	IN	IN	IN	OUT	OUT	N/A	OUT
SpDGAT2	**TM1**	29-49	29-49	26-45	32-51	—	29-43	25-47	—
	TM2	54-74	51-71	48-72	60-77	49-65	46-70	53-75	49-65
	TM3	115-135	—	119-137	115-134	115-137	119-133	—	115-137
	TM4	141-161	—	—	143-162	142-162	—	136-153	142-162
	TM5	217-237	—	—	217-236	217-237	—	—	217-237
	TM6	—	—	—	259-278	—	—	—	—
	Orientation	OUT	IN	IN	IN	OUT	OUT	N/A	OUT

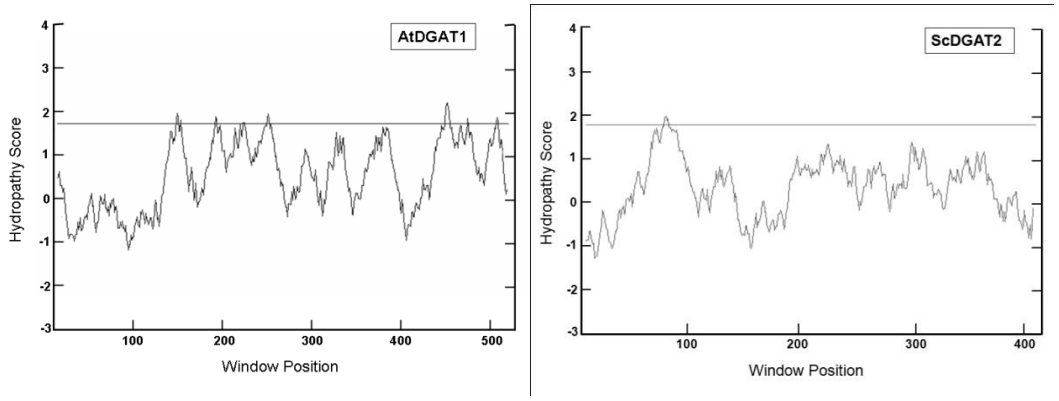
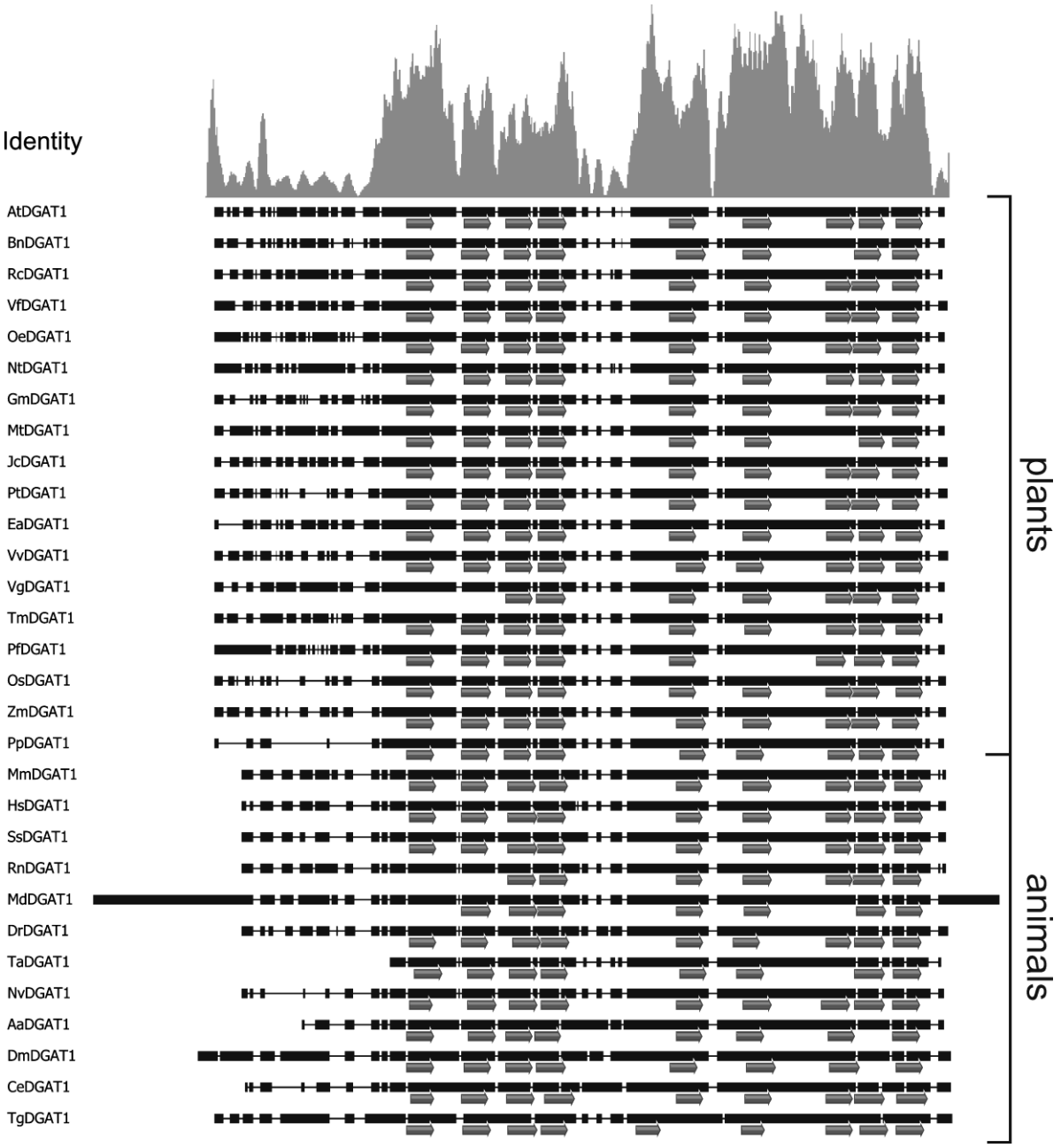


Figure 1: Kyte-Doolittle hydropathy plots of DGATs. Plots were generated by the method of Kyte and Doolittle (1982) using a window size of 19. Cut-off value (redline) is 1.8.

Figure 2: Alignment of transmembrane domains in DGAT1. The putative transmembrane domains of DGAT1 polypeptides from 12 animal and 18 plant organisms were predicted and the polypeptides were aligned. The identity of the alignment is graphed on the top using a window size of 6. The arrows denote the predicted transmembrane domains. The thick lines represent the sequence of each DGAT1 and the thin lines represent the gaps generated by the alignment. The picture was generated with Geneious Pro 4.6.0 and optimized manually. The transmembrane domains were predicted with Transmembrane Hidden Markov Model (TMHMM). Accession numbers for the DGAT1 polypeptides are: AtDGAT1, NM_127503; AaDGAT1, XP_001658299; BnDGAT1, AAD45536; CeDGAT1, CAB07399; DmDGAT1, AAL78365; DrDGAT1, NP_956024; EaDGAT1, AAV31083; GmDGAT1, AAS78662; HsDGAT1, NP_036211; JcDGAT1, ABB84383; MdDGAT1, XP_001371565; MmDGAT1, NP_034176; MtDGAT1, ABN09107; NtDGAT1, AAF19345; NvDGAT1, XP_001639351; OeDGAT1, AAS01606; OsDGAT1, BAD53762; PfDGAT1, AAG23696; PpDGAT1, XP_001770929; PtDGAT1, XP_002330510; RcDGAT1, AAR11479; RnDGAT1, BAC43739; SsDGAT1, NP_999216; TaDGAT1, XP_002112025; TgDGAT1, AAP94209; TmDGAT1, AAM03340; VfDGAT1, ABC94471; VgDGAT1, ABV21945; VvDGAT1, CAN80418; ZmDGAT1, ABV91586.


Identity
AtDGAT1
BnDGAT1
RcDGAT1
VfDGAT1
OeDGAT1
NtDGAT1
GmDGAT1
MtDGAT1
JcDGAT1
PtDGAT1
EaDGAT1
VvDGAT1
VgDGAT1
TmDGAT1
PfDGAT1
OsDGAT1
ZmDGAT1
PpDGAT1
MmDGAT1
HsDGAT1
SsDGAT1
RnDGAT1
MdDGAT1
DrDGAT1
TaDGAT1
NvDGAT1
AaDGAT1
DmDGAT1
CeDGAT1
TgDGAT1
plants
animals

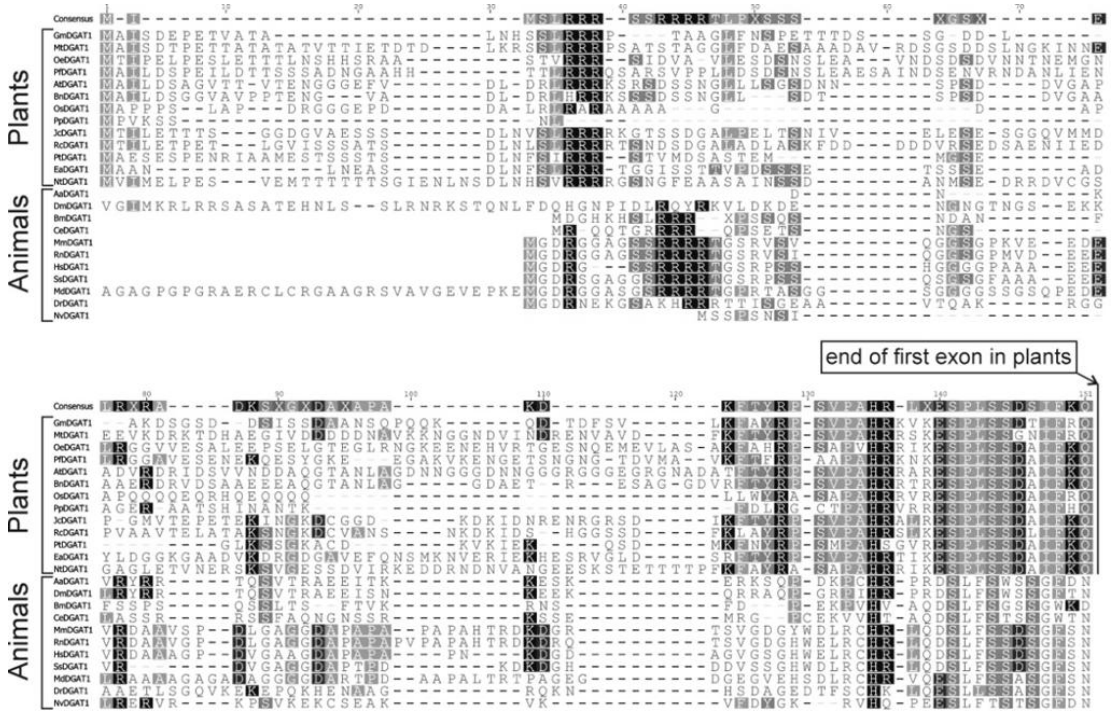
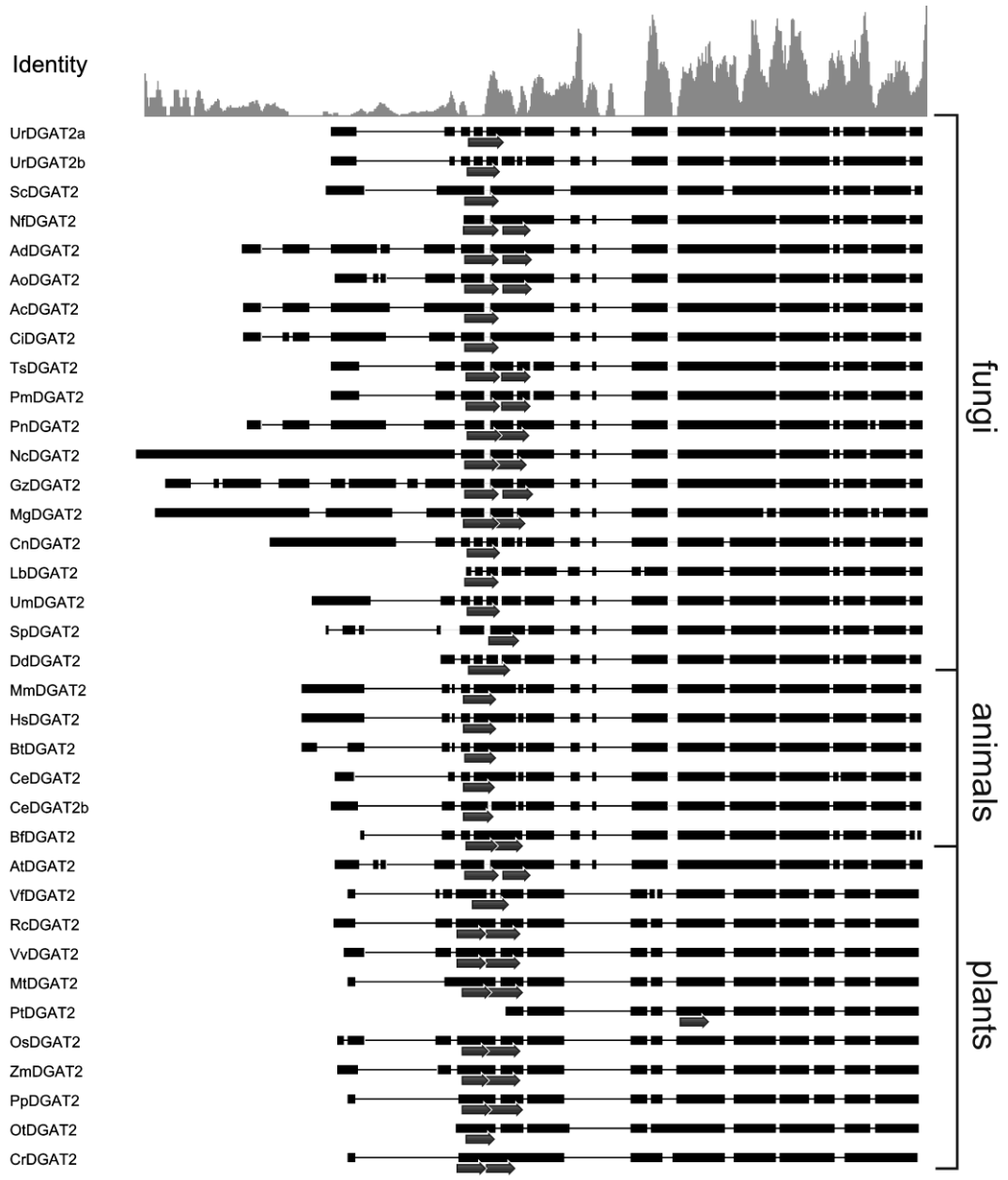


Figure 3: Alignment of the N-terminus polypeptide sequence of DGAT1 from plants and animals. Gray shades denote the polarity of blocks of conserved residues. The position corresponding to the end of the first exon in plants is indicated.

Figure 4: Alignment of transmembrane domains in DGAT2. The putative transmembrane domains of DGAT2 polypeptides from 18 fungi, 5 animals and 11 plant organisms were predicted and the polypeptides were aligned. The identity of the alignment is graphed on the top using a window size of 6. The arrows denote the predicted transmembrane domains. The thick lines represent the sequence of each DGAT2 polypeptide and the thin lines represent the gaps generated by the alignment. The picture was generated as described for DGAT1. Accession numbers for the DGAT2 polypeptides are: AcDGAT2, XP_001540241; AdDGAT2, XP_001273210; AnDGAT2, CAK46407; AoDGAT2, XP_001822244; AtDGAT2, NP_566952; BfDGAT2, XP_002208225; BtDGAT2, CAD58968; CeDGAT2, CAB04533; CeDGAT2b, AAB04969; CiDGAT2, XP_001240299; CnDGAT2, EAL20089; CrDGAT2, XP_001693189; DdDGAT2, XP_635762; GzDGAT2, XP_381525; HsDGAT2, AAK84176; LbDGAT2, EDR14458; MgDGAT2, XP_368741; MmDGAT2, AAK84175; MtDGAT2, ACJ84867; NcDGAT2, CAE76475; NfDGAT2, XP_001261291; OsDGAT2, NP_001057530; OtDGAT2, CAL58088; PmDGAT2, XP_002146410; PnDGAT2, EAT89076; PpDGAT2, XP_001777726; PtDGAT2, XP_002317635; RcDGAT2, AAY16324; ScDGAT2, NP_014888; SpDGAT2, XP_001713160; TsDGAT2, EED21737; UmDGAT2, XP_760084; UrDGAT2a, AAK84179; UrDGAT2b, AAK84180; VfDGAT2, ABC94474; VvDGAT2, CAO68497; ZmDGAT2, ACG38122;


Identity
UrDGAT2a
UrDGAT2b
ScDGAT2
NfDGAT2
AdDGAT2
AoDGAT2
AcDGAT2
CiDGAT2
TsDGAT2
PmDGAT2
PnDGAT2
NcDGAT2
GzDGAT2
MgDGAT2
CnDGAT2
LbDGAT2
UmDGAT2
SpDGAT2
DdDGAT2
MmDGAT2
HsDGAT2
BtDGAT2
CeDGAT2
CeDGAT2b
BfDGAT2
AtDGAT2
VfDGAT2
RcDGAT2
VvDGAT2
MtDGAT2
PtDGAT2
OsDGAT2
ZmDGAT2
PpDGAT2
OtDGAT2
CrDGAT2
fungi
animals
plants

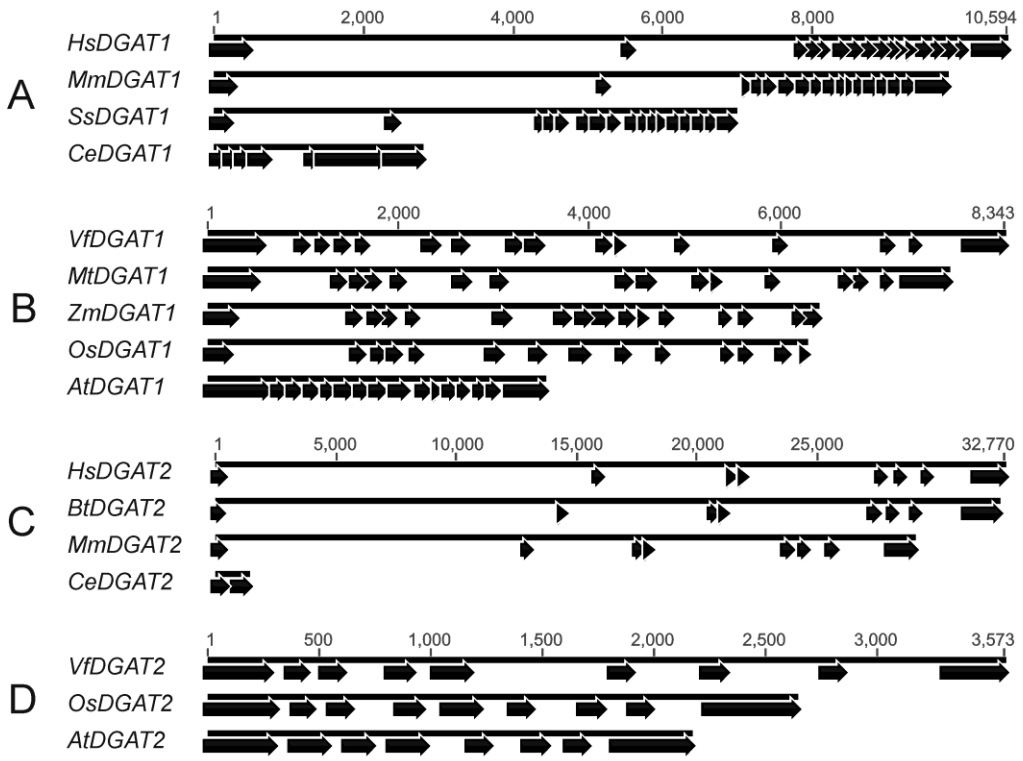


Figure 5: Architecture of DGAT genes. (A) *DGAT1* from animals, (B) *DGAT1* from plants, (C) *DGAT2* from animals and (D) *DGAT2* from plants. The genomic sequences of each DGAT are represented by black bars. The arrows correspond to the regions comprising the coding region of the mRNA. The numbers correspond to the nucleotide positions. Accession numbers are: *HsDGAT1*, AC_000140.1; *MmDGAT1*, NC_000081.5; *SsDGAT1*, AY116586.1; *CeDGAT1*, NC_003283.9; *VfDGAT1*, DQ356679.1; *MtDGAT1*, AC174465.2; *ZmDGAT1*, AM433916.2; *OsDGAT1*, AP008212.1; *AtDGAT1*, NC_003071.4; *HsDGAT2*, NC_000011.8; *BtDGAT2*, NC_007313.3; *MmDGAT2*, NC_000073.5; *CeDGAT2*, Z81557.1; *VfDGAT2*, DQ356681.1; *OsDGAT2*, AP004757.3 and *AtDGAT2*, NC_003074.5.

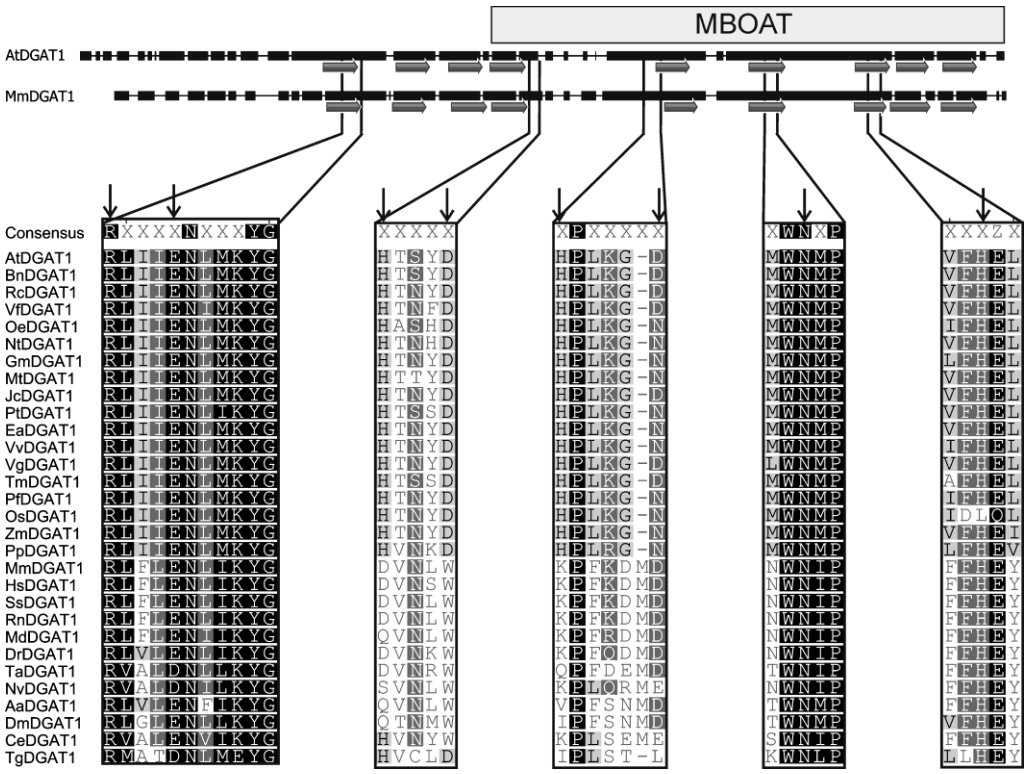


Figure 6: Alignment of putative active sites in DGAT1. A scheme MmDGAT1and AtDGAT1 is described on the top with the position of the MBOAT motif. The arrows in this scheme represent the predicted transmembrane domains. The thick lines represent the sequence of each DGAT1 polypeptide and the thin lines represent the gaps generated though the alignment previously shown. The vertical boxes contain the amino acid sequences for different DGATs indicated on the left. The arrows on these boxes indicate the position of conserved residues discussed on the text. Accession numbers for the DGAT polypeptides are the same as in the Figure 2.

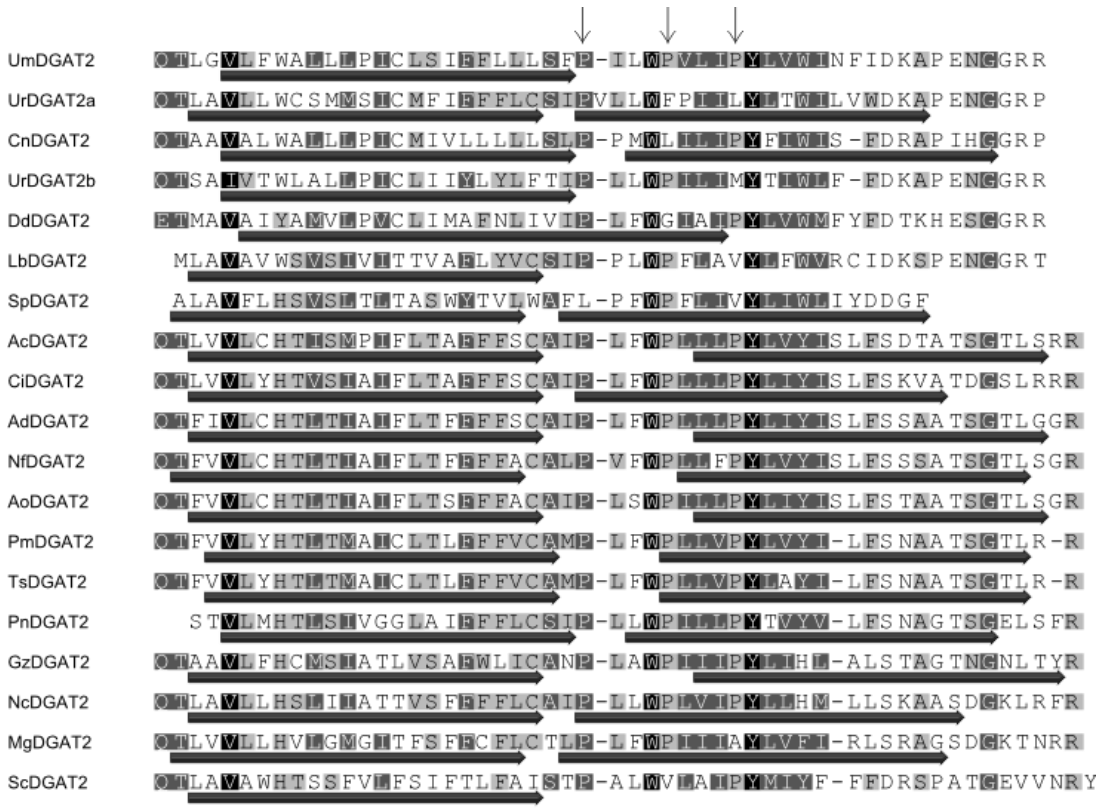


Figure 7: Alignment of predicted transmembrane domain region from fungi DGAT2. Residues are highlighted in different shades of gray to black according to their similarity. The arrows denote the position of the predicted transmembrane domains. The positions of prolines conserved in many sequences are denoted by the arrows on the top. The alignment parameters and the accession numbers are the same as previously indicated in Figure 4.

References

Abell, B.M., Hahn, M., Holbrook, L.A., and Moloney, M.M. (2004). Membrane topology and sequence requirements for oil body targeting of oleosin. *Plant Journa*l 37, 461-470.

Bouvier-Nave, P., Benveniste, P., Noiriel, A., and Schaller, H. (2000a). Expression in yeast of an acyl-CoA:diacylglycerol acyltransferase cDNA from Caenorhabditis elegans. *Biochemical Society Transactions* 28, 692-695.

Bouvier-Nave, P., Benveniste, P., Oelkers, P., Sturley, S.L., and Schaller, H. (2000b). Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *European Journal of Biochemistry* 267, 85-96.

Browse, J., and Somerville, C. (1991). Glycerolipid synthesis-biochemistry and regulation. *Annual Review of Plant Physiology and Plant Molecular Biology* 42, 467-506.

Burgal, J., Shockey, J., Lu, C.F., Dyer, J., Larson, T., Graham, I., and Browse, J. (2008). Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnology Journal* 6, 819-831.

Cahoon, E.B., Shockey, J.M., Dietrich, C.R., Gidda, S.K., Mullen, R.T., and Dyer, J.M. (2007). Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Current Opinion Plant Biology*. 10, 236-244.

Cao, Y.Z., and Huang, A.H.C. (1986). Diacylglycerol acyltransferase in maturing oil seeds of maize and other species. *Plant Physiology* 82, 813-820.

Carrasco, S., and Merida, I. (2007). Diacylglycerol, when simplicity becomes complex. *Trends in Biochemical Sciences* 32, 27-36.

Cases, S., Stone, S.J., Zhou, P., Yen, E., Tow, B., Lardizabal, K.D., Voelker, T., and Farese, R.V. (2001). Cloning of DGAT2, a second mammalian diacylglycerol

acyltransferase, and related family members. *Journal of Biological Chemistry* 276, 38870-38876.

Cases, S., Smith, S.J., Zheng, Y.W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusis, A.J., Erickson, S.K., and Farese, R.V. (1998). Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 95, 13018-13023.

Chamoun, Z., Mann, R.K., Nellen, D., von Kessler, D.P., Bellotto, M., Beachy, P.A., and Basler, K. (2001). Skinny Hedgehog, an acyltransferase required for palmitoylation and activity of the Hedgehog signal. *Science* 293, 2080-2084.

Chang, C.C.Y., Huh, H.Y., Cadigan, K.M., and Chang, T.Y. (1993). Molecular cloning and functional expression of human acyl-Coenzyme A:cholesterol acyltransferase cDNA in mutant Chinese hamster ovary cells. *Journal of Biological Chemistry* 268, 20747-20755.

Cheng, D., Meegalla, R.L., He, B.K., Cromley, D.A., Billheimer, J.T., and Young, P.R. (2001). Human acyl-CoA:diacylglycerol acyltransferase is a tetrameric protein. *Biochemical Journal* 359, 707-714.

Cheng, D., Nelson, T.C., Chen, J., Walker, S.G., Wardwell-Swanson, J., Meegalla, R., Taub, R., Billheimer, J.T., Ramaker, M., and Feder, J.N. (2003). Identification of acyl coenzyme A:Monoacylglycerol acyltransferase 3, an intestinal specific enzyme implicated in dietary fat absorption. *Journal of Biological Chemistry* 278, 13611-13614.

Coleman, R.A. (1992). Diacylglycerol acyltransferase and monoacylglycerol acyltransferase from liver and intestine. *Methods in Enzymology* 209, 98-104.

Daniel, J., Deb, C., Dubey, V.S., Sirakova, T.D., Abomoelak, B., Morbidoni, H.R., and Kolattukudy, P.E. (2004). Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture. *Journal of Bacteriology* 186, 5017-5030.

Dyer, J.M., and Mullen, R.T. (2008). Engineering plant oils as high-value industrial feedstocks for biorefining: the need for underpinning cell biology research. *Physiologia Plantarum* 132, 11-22.

Foroud, N.A. (2005). Probing the membrane topology of a diacylglycerol acyltransferase type I from Brassica napus. In Department of Chemistry and Biochemistry (Lethbridge: University of Lethbridge), pp. 194.

Gangar, A., Karande, A.A., and Rajasekharan, R. (2001). Isolation and localization of a cytosolic 10 S triacylglycerol biosynthetic multienzyme complex from oleaginous yeast. *Journal of Biological Chemistry* 276, 10290-10298.

Guo, Z.M., Cromley, D., Billheimer, J.T., and Sturley, S.L. (2001). Identification of potential substrate-binding sites in yeast and human acyl-CoA sterol acyltransferases by mutagenesis of conserved sequences. *Journal of Lipid Research* 42, 1282-1291.

Guo, Z.Y., Lin, S., Heinen, J.A., Chang, C.C.Y., and Chang, T.Y. (2005). The active site His-460 of human acyl-coenzyme A : cholesterol acyltransferase 1 resides in a hitherto undisclosed transmembrane domain. *Journal of Biological Chemistry* 280, 37814-37826.

He, X.H., Turner, C., Chen, G.Q., Lin, J.T., and McKeon, T.A. (2004). Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. *Lipids* 39, 311-318.

Heath, R.J., and Rock, C.O. (1998). A conserved histidine is essential for glycerolipid acyltransferase catalysis. *Journal of Bacteriology* 180, 1425-1430.

Higy, M., Junne, T., and Spiess, M. (2004). Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* 43, 12716-12722.

Hildebrand, D.F., Li, R., Hatanaka, T., and G.Stacey. (2008). Genomics of soybean oil traits. In Genetics and Genomics of Soybean (New York: Springer), pp. 185-209.

Hobbs, D.H., Lu, C.F., and Hills, M.J. (1999). Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. *FEBS Letters* 452, 145-149.

Hofmann, K. (2000). A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. *Trends in Biochemical Sciences* 25, 111-112.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.

Ikeda, M., Arai, M., Lao, D.M., and Shimizu, T. (2002). Transmembrane topology prediction methods: a re-assessment and improvement by a consensus method using a dataset of experimentally-characterized transmembrane topologies. *In Silico Biology* 2, 19-33.

Inokoshi, J., Kawamoto, K., Takagi, Y., Matsuhama, M., Omura, S., and Tomoda, H. (2009). Expression of two human acyl-CoA:diacylglycerol acyltransferase isozymes in yeast and selectivity of microbial inhibitors toward the isozymes. *Journal of Antibiotics* 62, 51-54.

Jako, C., Kumar, A., Wei, Y.D., Zou, J.T., Barton, D.L., Giblin, E.M., Covello, P.S., and Taylor, D.C. (2001). Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiology* 126, 861-874.

Kalscheuer, R., and Steinbuchel, A. (2003). A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in Acinetobacter calcoaceticus ADP1. *Journal of Biological Chemistry* 278, 8075-8082.

Katavic, V., Reed, D.W., Taylor, D.C., Giblin, E.M., Barton, D.L., Zou, J.T., MacKenzie, S.L., Covello, P.S., and Kunst, L. (1995). Alteration of seed fatty-acid composition by an ethyl methanesulfonate-induced mutation in Arabidopsis thaliana affecting diacylglycerol acyltransferase activity. *Plant Physiology* 108, 399-409.

Kaup, M.T., Froese, C.D., and Thompson, J.E. (2002). A role for diacylglycerol acyltransferase during leaf senescence. *Plant Physiology* 129, 1616-1626.

Kroon, J.T., Wei, Y.D., Simon, W.J., and Slabas, A.R. (2006). Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. *Phytochemistry* 67, 2541-2549.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157, 105-132.

Lacey, D.J., and Hills, M.J. (1996). Heterogeneity of the endoplasmic reticulum with respect to lipid synthesis in developing seeds of Brassica napus L. *Planta* 199, 545-551.

Lardizabal, K., Effertz, R., Levering, C., Mai, J., Pedroso, M.C., Jury, T., Aasen, E., Gruys, K., and Bennett, K. (2008). Expression of Umbelopsis ramanniana DGAT2A in seed increases oil in soybean. *Plant Physiology* 148, 89-96.

Lardizabal, K.D., Mai, J.T., Wagner, N.W., Wyrick, A., Voelker, T., and Hawkins, D.J. (2001). DGAT2 is a new diacylglycerol acyltransferase gene family - Purification, cloning, and expression in insect cells of two polypeptides from Mortierella ramanniana with diacylglycerol acyltransferase activity. *Journal of Biological Chemistry* 276, 38862-38869.

Leber, R., Zinser, E., Zellnig, G., Paltauf, F., and Daum, G. (1994). Characterization of lipid particles of the yeast, *Saccharomyces cerevisiae*. *Yeast* 10, 1421-1428.

Lewin, T.M., Wang, P., and Coleman, R.A. (1999). Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate acyltransferase reaction. *Biochemistry* 38, 5764-5771.

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., Jetter, R., and Kunst, L. (2008). Identification of the wax ester synthase/acyl-coenzyme A:Diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in *Arabidopsis*. *Plant Physiology* 148, 97-107.

Liang, J.J., Oelkers, P., Guo, C.Y., Chu, P.C., Dixon, J.L., Ginsberg, H.N., and Sturley, S.L. (2004). Overexpression of human diacylglycerol acyltransferase 1, acyl-CoA:cholesterol acyltransferase 1, or acyl-CoA : cholesterol acyltransferase 2

stimulates secretion of apolipoprotein B-containing lipoproteins in McA-RH7777 cells. *Journal of Biological Chemistry* 279, 44938-44944.

Lu, C.F.L., de Noyer, S.B., Hobbs, D.H., Kang, J.L., Wen, Y.C., Krachtus, D., and Hills, M.J. (2003). Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in triacylglycerol biosynthesis, in *Arabidopsis thaliana*. *Plant Molecular Biology* 52, 31-41.

Lung, S.C., and Weselake, R.J. (2006). Diacylglycerol acyltransferase: A key mediator of plant triacylglycerol synthesis. *Lipids* 41, 1073-1088.

Martin, B.A., and Wilson, R.F. (1984). Subcellular localization of triacylglycerol synthesis in spinach leaves. *Lipids* 19, 117-121.

Moller S., Croning M.D.R. and Apweiler R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646-653.

Milcamps, A., Tumaney, A.W., Paddock, T., Pan, D.A., Ohlrogge, J., and Pollard, M. (2005). Isolation of a gene encoding a 1,2-diacylglycerol-sn-acetyl-CoA acetyltransferase from developing seeds of Euonymus alatus. *Journal of Biological Chemistry* 280, 5370-5377.

Natter, K., Leitner, P., Faschinger, A., Wolinski, H., McCraith, S., Fields, S., and Kohlwein, S.D. (2005). The spatial organization of lipid synthesis in the yeast Saccharomyces cerevisiae derived from large scale green fluorescent protein tagging and high resolution microscopy. *Molecular & Cellular Proteomics* 4, 662-672.

Nykiforuk, C.L., Furukawa-Stoffer, T.L., Huff, P.W., Sarna, M., Laroche, A., Moloney, M.M., and Weselake, R.J. (2002). Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of Brassica napus and sucrose-mediated induction of enzyme biosynthesis. *Biochimica et Biophysica Acta-Molecular and Cell Biology of Lipids* 1580, 95-109.

Palacpac, N.M.Q., Hiramine, Y., Seto, S., Hiramatsu, R., Horii, T., and Mitamura, T. (2004). Evidence that Plasmodium falciparum diacylglycerol acyltransferase is essential for intraerythrocytic proliferation. *Biochemical and Biophysical Research Communications* 321, 1062-1068.

Perry, H.J., and Harwood, J.L. (1993). Changes in the lipid content of developing seeds of Brassica napus. *Phytochemistry* 32, 1411-1415.

Persson, B. (2006). Bioinformatics in membrane protein analysis. In Structual genomics of membrane proteins, Lundstrom K.H., ed (Florida: CRC press), pp. 5-20.

Quittnat, F., Nishikawa, Y., Stedman, T.T., Voelker, D.R., Choi, J.Y., Zahn, M.M., Murphy, R.C., Barkley, R.M., Pypaert, M., Joiner, K.A., and Coppens, I. (2004). On the biogenesis of lipid bodies in ancient eukaryotes: synthesis of triacylglycerols by a Toxoplasma DGAT1-related enzyme. *Molecular and Biochemical Parasitology* 138, 107-122.

Saha, S., Enugutti, B., Rajakumari, S., and Rajasekharan, R. (2006). Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiology* 141, 1533-1543.

Schutze, M.P., Peterson, P.A., and Jackson, M.R. (1994). An N-Terminal double-arginine motif maintains type-II membrane-proteins in the endoplasmic-reticulum. *Embo Journal* 13, 1696-1705.

Settlage, S.B., Wilson, R.F., and Kwanyuen, P. (1995). Localization of diacylglycerol acyltransferase to oil body associated endoplasmic reticulum. *Plant Physiology and Biochemistry* 33, 399-407.

Shockey, J.M., Gidda, S.K., Chapital, D.C., Kuan, J.C., Dhanoa, P.K., Bland, J.M., Rothstein, S.J., Mullen, R.T., and Dyer, J.M. (2006). Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* 18, 2294-2313.

Siloto, R.M.P., Madhavji, M., Wiehler, W.B., Burton, T.L., Boora, P.S., Laroche, A., and Weselake, R.J. (2008). An N-terminal fragment of mouse DGAT1 binds different acyl-CoAs with varying affinity. *Biochemical and Biophysical Research Communications 3*73, 350-354.

Siloto R.M.P., Truksa M., Brownfield D., Good A.G. and Weselake R.J. (2009) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and

characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiology and Biochemistry* 47, 456-461

Sorger, D., and Daum, G. (2002). Synthesis of triacylglycerols by the acyl-coenzyme A: diacyl-glycerol acyltransferase Dga1p in lipid particles of the yeast *Saccharomyces cerevisiae*. *Journal of Bacteriology* 184, 519-524.

Stone, S.J., Levin, M.C., and Farese, R.V. (2006). Membrane topology and identification of key functional amino acid residues of murine acyl-CoA: diacylglycerol acyltransferase-2. *Journal of Biological Chemistry* 281, 40273-40282.

Stone, S.J., Levin, M.C., Zhou, P., Han, J.Y., Walther, T.C., and Farese, R.V. (2009). The endoplasmic reticulum enzyme DGAT2 is found in mitochondria-associated membranes and has a mitochondrial targeting signal that promotes its association with mitochondria. *Journal of Biological Chemistry* 284, 5352-5361.

Teasdale, R.D., and Jackson, M.R. (1996). Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annual Review of Cell and Developmental Biology 12*, 27-54.

Thuille, N., Heit, I., Fresser, F., Krumbock, N., Bauer, B., Leuthaeusser, S., Darnmeier, S., Graham, C., Copeland, T.D., Shaw, S., and Baier, G. (2005). Critical role of novel Thr-219 autophosphorylation for the cellular function of PKC theta in T lymphocytes. *Embo Journal* 24, 3869-3880.

Turkish, A.R., Henneberry, A.L., Cromley, D., Padamsee, M., Oelkers, P., Bazzi, H., Christiano, A.M., Billheimer, J.T., and Sturley, S.L. (2005). Identification of two novel human acyl-CoA wax alcohol acyltransferases: members of the diacylglycerol acyltransferase 2 (DGAT2) gene superfamily. *Journal of Biological Chemistry* 280, 14755-14764.

Tzen, J.T.C., Lie, G.C., and Huang, A.H.C. (1992). Characterization of the charged components and their topology on the surface of plant seed oil bodies. *Journal of Biological Chemistry* 267, 15626-15634.

Tzen, J.T.C., Cao, Y.Z., Laurent, P., Ratnayake, C., and Huang, A.H.C. (1993). Lipids, proteins, and structure of seed oil bodies from diverse species. *Plant Physiology* 101, 267-276.

Weiss, S.B., and Kennedy, E.P. (1956). The enzymatic synthesis of triglycerides. *Journal of the American Chemical Society* 78, 3550-3550.

Weiss, S.B., Kennedy, E.P., and Kiyasu, J.Y. (1960). Enzymatic synthesis of triglycerides. *Journal of Biological Chemistry* 235, 40-44.

Weselake, R.J. (2005). Storage lipids. In *Plant Lipids: Biology, Utilization and Manipulation*, Murphy. D.J., ed (Oxford: Blackwell publishing), pp. 162-225.

Weselake, R.J, Kuo, T.M., and Gardner, H.W. (2002). Biochemistry and biotechnology of TAG accumulation in Plants. In *Lipid Biotechnology* (Peoria, IL: Marcel Dekker), pp. 27-56.

Weselake, R.J., Madhavji, M., Szarka, S.J., Patterson, N.A., Wiehler, W.B., Nykiforuk, C.L., Burton, T.L., Boora, P.S., Mosimann, S.C., Foroud, N.A., Thibault, B.J., Moloney, M.M., Laroche, A., and Furukawa-Stoffer, T.L. (2006). Acyl-CoA-binding and self-associating properties of a recombinant 13.3 kDa N-terminal fragment of diacylglycerol acyltransferase-1 from oilseed rape. *BMC Biochemistry*. 7, 24.

Weselake, R.J., Nykiforuk, C.L., Laroche, A., Patterson, N.A., Wiehler, W.B., Szarka, S.J., Moloney, M.M., Tari, L.W., and Derekh, U. (2000). Expression and properties of diacylglycerol acyltransferase from cell-suspension cultures of oilseed rape. *Biochemical Society Transactions* 28, 684-686.

Weselake, R.J., Shah, S., Tang, M.G., Quant, P.A., Snyder, C.L., Furukawa-Stoffer, T.L., Zhu, W.M., Taylor, D.C., Zou, J.T., Kumar, A., Hall, L., Laroche, A., Rakow, G., Raney, P., Moloney, M.M., and Harwood, J.L. (2008). Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. *Journal of Experimental Botany* 59, 3543-3549.

Xu, J.Y., Francis, T., Mietkiewska, E., Giblin, E.M., Barton, D.L., Zhang, Y., Zhang, M., and Taylor, D.C. (2008). Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed

mutagenesis to modify enzyme activity and oil content. *Plant Biotechnology Journal* 6, 799-818.

Yen, C.L., Stone, S.J., Koliwad, S., Harris, C., and Farese, R.V. (2008). DGAT enzymes and triacylglycerol biosynthesis. *Journal of Lipid Research* 49, 2283-2301.

Yen, C.L.E., and Farese, R.V. (2003). MGAT2, a monoacylglycerol acyltransferase expressed in the small intestine. *Journal of Biological Chemistry* 278, 18532-18537.

Yen, C.L.E., Monetti, M., Burri, B.J., and Farese, R.V. (2005). The triacylglycerol synthesis enzyme DGAT1 also catalyzes the synthesis of diacylglycerols, waxes, and retinyl esters. *Journal of Lipid Research* 46, 1502-1511.

Yen, C.L.E., Stone, S.J., Cases, S., Zhou, P., and Farese, R.V. (2002). Identification of a gene encoding MGAT1, a monoacylglycerol acyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 99, 8512-8517.

Yu, C.J., Chen, J., Lin, S., Chang, C.C.Y., and Chang, T.Y. (1999). Human acyl-CoA:cholesterol acyltransferase-1 is a homotetrameric enzyme in intact cells and in vitro. *Journal of Biological Chemistry* 274, 36139-36145.

Yu, C.J., Zhang, Y., Lu, X.H., Chen, J., Chang, C.C.Y., and Chang, T.Y. (2002). Role of the N-terminal hydrophilic domain of acyl-coenzyme A:cholesterol acyltransferase I on the enzyme's quaternary structure and catalytic efficiency. *Biochemistry* 41, 3762-3769.

Yu, K., Li, R., Hatanaka, T., and Hildebrand, D. (2008). Cloning and functional analysis of two type 1 diacylglycerol acyltransferases from Vernonia galamensis. *Phytochemistry* 69, 1119-1127.

Zheng, P., Allen, W.B., Roesler, K., Williams, M.E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D., Solawetz, W., Bhattramakki, D., Llaca, V., Deschamps, S., Zhong, G.Y., Tarczynski, M.C., and Shen, B. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nature Genetics* 40, 367-372.

Zou, J.T., Wei, Y.D., Jako, C., Kumar, A., Selvaraj, G., and Taylor, D.C. (1999). The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *Plant Journal* 19, 645-653.

A version of Appendix.2 has been published. Snyder, C. L. Yurchenko O. P., Siloto, R. M. P, Chen, X., Liu, Q, Mietkiewska, E. and Weselake, R. J. 2009. New biotechnology, 26, 11-16.

Appendix. 2 Acyltransferase Action in the Modification of Seed Oil Biosynthesis

(Reprinted with permission of Elsevier Science Ltd.)

I was responsible for writing sections related to acyl-CoA:diacylglycerol acyltransferase 2 (DGAT2)

Abstract:

Seed oils represent a major source of dietary lipid and an increasingly valuable feedstock for industrial applications. There have been a number of attempts to modify seed oil content and composition through biotechnological approaches, resulting in the identification of several bottlenecks limiting the accumulation of unusual fatty acids in storage lipids of oilseed crops. It has been suggested that the substrate preferences of endogenous acyltransferases play an important role in the utilization of unusual fatty acids in transgenic oilseeds, and there is increasing evidence that mechanisms of "acyl-editing" via phospholipids are also involved in substrate trafficking and utilization. In this review, we will examine acyltransferase substrate specificity and selectivity in the context of designing strategies to maximize the accumulation of unusual fatty acids using biotechnological approaches.

Introduction

Soaring global demand for plant oils for edible and industrial applications has sparked a considerable amount of research focused on biotechnological approaches to oil modification (Drexler et al., 2003) (**Table 1)**. Despite steady progress in the cloning and characterization of genes involved in plant storage lipid metabolism, attempts to produce unusual fatty acids at commercially viable levels in oilseed crops have largely fallen short. A number of potential bottlenecks have been identified, including substrate dichotomy between acyl-CoA and phospholipid pools and the inability of endogenous acyltransferases to efficiently incorporate the desired acyl groups into storage lipids (Larson et al., 2002), reviewed by (Drexler et al., 2003; Cahoon et al., 2007). In this review, we will discuss the role of acyltransferase substrate preferences in the modification of seed oils and discuss strategies for overcoming these obstacles through biotechnological approaches.

Acyltransferases involved in plant storage lipid biosynthesis

The pathways involved in plant storage lipid biosynthesis have been recently reviewed in detail elsewhere (Harwood, 2005; Weselake, 2005) and are summarized in **Figure 1**. For the purposes of this review, we will consider acyltransferases involved in two pathways leading to triacylglycerol (TAG). The classical *sn*-glycerol-3-phosphate or Kennedy pathway involves the sequential acyl-CoA dependent acylation of *sn*-glycerol-

3-phosphate catalyzed by *sn*-glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAAT), and diacylglycerol acyltransferase (DGAT), respectively, with phosphatidic acid phosphatase catalyzing the dephosphorylation of PA prior to the final acylation. In a second pathway, acyl groups can be channeled into phosphatidylcholine (PC) via the activity of lysophosphatidylcholine acyltransferase (LPCAT), and then subsequently transferred from PC to TAG via the activity of phospholipid:diacylglycerol acyltransferase (PDAT) (Dahlqvist et al., 2000; Stahl et al., 2004). This latter pathway represents one of several possible mechanisms for "acyl editing", permitting acyl groups which are modified at the level of PC (i.e. by desaturation of hydroxylation) to be channeled into storage lipids. Understanding these acyl-editing mechanisms and the substrate preferences of the enzymes involved is important for developing strategies for overcoming the problem of substrate dichotomy in the accumulation of unusual fatty acids in transgenic plants.

Substrate specificity properties of the Kennedy Pathway enzymes

sn-glycerol-3-phosphate acyltransferase (GPAT)

Of the three known types of plant GPATs (plastidial, mitochondrial, and endoplasmic reticulum), only the ER-bound GPAT is known to be involved in the Kennedy pathway. To date, five *Arabidopsis* genes encoding ER-bound GPAT have been reported (Beisson et al., 2007; Li et al., 2007; Zheng et al., 2003), mostly with roles

in suberin and cutin synthesis. The role of specific GPAT genes in TAG synthesis has not been clarified to date. While there are few published reports on the substrate specificity of ER-bound GPAT enzymes, positional analyses of several plant oils have suggested that plant GPATs have a relatively broad specificity, utilizing saturated, monounsaturated, and polyunsaturated moieties (Brokerhoff et al., 1966; Christie et al., 1991; Lisa et al., 2008). Dutta et al. (1992) demonstrated that microsomal GPATs from safflower and rapeseed utilized oleoyl-CoA ($18:1^{cis\Delta 9}$) more efficiently than petroselinoyl-CoA ($18:1^{cis\Delta 6}$). Subsequent studies of transgenic *Arabidopsis*, however, showed that petroselinic acid was incorporated into TAG as effectively as oleic acid (Suh et al., 2002). In *B. napus* engineered to produce medium chain saturated fatty acids, up to 80% of the acyl groups at the *sn*-1 position were medium chain saturated fatty acids (Voelker et al., 2001; Voelker et al., 1996; Wiberg et al., 2000), whereas *B. napus* normally contains about 7% saturated fatty acids, mainly occurring at the *sn*-1 position. These observations support the view that GPAT may not exhibit a strong substrate preference *in vivo*, rather, acyl-CoA availability may play a more pivotal role in determining the acyl composition at the *sn*-1 position of TAG (Voelker et al., 2001).

Lysophosphatidic acid acyltransferase (LPAAT)

Microsomal LPAAT has the highest substrate stringency of the three Kennedy pathway acyltransferases (Laurent et al., 1992). Plants producing unusual FAs often have

microsomal LPAAT activities with enhanced preferences for these FAs (Bafor et al., 1990; Bernerth etl al., 1990; Cao et al., 1990; Davies et al., 1995; Lohden et al., 1992; Oo et al., 1989; Sun et al., 1988;20-26). In addition to its role in storage lipid assembly, LPAAT also plays a role in the *de novo* synthesis of PC and this may be reflected in the acyl-donor specificity of LPAATs containing high levels of polyunsaturated fatty acids. For example, it has been shown that flax microsomal LPAAT prefers 18:2$^{cis\Delta 9,12}$-CoA > 18:1$^{cis\Delta 9}$-CoA > 18:3$^{cis\Delta 9,12,15}$-CoA (Sorensen et al., 2005). This specificity is consistent with a possible role for LPAAT in channeling substrates toward the *sn*-2 position of PC for further desaturation to α-18:3, which constitutes more than 60% of the FA in mature flax seed. Brown et al. (2002) have shown that microsomal LPAATs from several plant species (including *B. napus*) display little selectivity between 18:1- and 18:2-CoA donor substrates, whereas 16:0- and 18:0-CoA are generally discriminated against. Such discrimination against saturated acyl groups in the *sn*-2 position of TAG is relatively common among major oilseed crops (11,12), and hampered early progress toward engineering laurate (12:0) accumulation in transgenic *B. napus* to produce cocoa butter-like oils (Frentzen, 1993; Sun et al., 1988). Incorporation of 12:0 at the *sn*-2 position of TAG was eventually accomplished by expression of a microsomal laurate-specific LPAAT from coconut (*Cocos nucifera*) in *B. napus* expressing a California bay laurel (*Umbellularia californica*) 12:0-ACP thioesterase (Knutzon et al., 1999). Similarly,

LPAATs with enhanced specificity for erucoyl-CoA (22:1$^{cis\Delta 13}$) were expressed in *B. napus*, resulting in substantial quantities of 22:1 at the *sn*-2 position of TAG (Brough et al., 1996; Katavic et al., 2001; Lassner et al., 1995; 31-35; Taylor et al., 2002; Zou et al., 1997), overcoming the tendency for *B. napus* LPAAT to exclude 22:1 from this position (Bernerth et al., 1990; Taylor et al., 1991). Oils with a high 22:1 content can serve as feedstocks in the preparation of detergents, plasticizers, surface coatings and slip-promoting/anti-blocking agents with trierucin serving as a high temperature lubricant (Princen et al., 1984).

Diacylglycerol acyltransferase (DGAT)

DGAT has the lowest specific activity of the Kennedy pathway enzymes and is regarded as a "bottleneck" in TAG synthesis (Perry et al., 1993). As such, DGAT has been one of the more extensively studied enzymes both in terms of its substrate preferences and for the potential to increase seed oil content through DGAT over-expression. Early biochemical studies evaluated DGAT substrate preference through *in vitro* assays using microsomal fractions of developing seeds. In safflower, it was shown that DGAT can utilize a broad range of molecular species of DAG and acyl-CoA (Ichihara et al., 1982, 1988). Other studies have indicated that in species such as *B. napus, Tropaeolum majus* (Nasturtium), *Ricinus communis* (Castor bean) and *Cuphea lanceolata*, endogenous DGATs exhibit preference for substrates with defined acyl

chains (Lohden et al., 1992; Lung et al., 2006; He et al., 2004; Martin et al., 1983; Taylor et al., 1991; Taylor et al., 1992; Vogel et al., 1996; Yu et al., 2006).

It is important to note that early *in vitro* studies on DGAT substrate preference did not distinguish between type-1 DGATs (DGAT1) and the more recently identified type-2 DGATs (DGAT2). The discovery of genes encoding these two non-homologous DGAT polypeptides has allowed a more refined dissection of TAG biosynthesis through the study of individual genes in recombinant systems. In *Vernicia fordii* (tung tree), DGAT1 and DGAT2 display minor differences in substrate selectivity as determined by *in vitro* assays. *In vivo* selectivity, evaluated through a yeast expression system, has revealed that DGAT2 is able to synthesize higher amounts of trieleostearin, an abundant triacylglycerol found in tung tree (Shockey et al., 2006). Moreover, it has been shown that these two DGATs are present in distinctive ER domains, corroborating the hypothesis that they do not have overlapping functions. Similarly, in castor bean, DGAT2 is selective for substrates with ricinoleic acid and is responsible for the accumulation of triricinolein (Burgal et al., 2008). Burgal et al (2008) reported that co-expression of fatty acid hydroxylase (*FAH12*) and *DGAT2* from castor bean resulted in up to 30% hydroxy fatty acids, nearly double that attainable from expression of *FAH12* alone. Such studies highlight the importance of using acyltransferases with appropriate substrate preferences for maximizing the accumulation of unusual fatty acids in transgenic plants.

Although these results suggest that DGAT2 may be more important than DGAT1 in plants, the actual contribution of each isoform to seed oil accumulation is still uncertain as only DGAT1 knockouts have been characterized in plants. In *Arabidopsis*, an EMS-induced mutation of DGAT1 (AS11) resulted in a 15% decrease in seed lipids (Katavic et al., 1995). In maize, a high-oil quantitative trait locus *qHO6* was shown to be caused by a single residue insertion in DGAT1 polypeptide and $qHO6^{-/-}$ maize lines possessed lower oil content (Zheng et al., 2008). In both *Arabidopsis* and maize a decrease of DGAT1 activity also affected the fatty acid composition, resulting in an increase of α-linolenic acid content. This phenotype could be reversed by restoring the DGAT1 activity (Jako et al., 2001; Zheng et al., 2008), suggesting that in these species DGAT1 might also influence certain acyl channeling to TAG.

Both *DGAT1* and *DGAT2* have been successfully over-expressed in crops as a means of increasing seed oil content. *B. napus* expressing *DGAT1* (Taylor et al., 2009; Weselake et al., 2008) and soybean expressing *DGAT2* (Lardizabal et al., 2008) both exhibited similar gains in oil content when tested under field conditions, suggesting that both enzymes may play a critical role in TAG accumulation despite differences in substrate preferences. The precise contribution of DGAT1 and DGAT2 to overall TAG content and composition remains an open area of investigation. Further biotechnological applications of DGAT in oilseed metabolic engineering will likely require a better

understanding of the mechanisms governing the enzyme activity. Molecular genetics approaches such as site directed mutagenesis and directed evolution could potentially be used to increase of the overall catalytic efficiency and modify substrate selectivity(Siloto et al., 2009).

The contribution of LPCAT and PDAT to acyl-trafficking via PC

It has been shown that majority of the fatty acids synthesised in plastids enter PC through acyl editing of PC rather than through *de novo* synthesis (Bates et al., 2007; Williams et al., 2000). These observations are supported by experiments demonstrating that LPCAT is one of the most active acyltransferases in seed oil biosynthesis (Ichihara et al., 1995). In soybean, safflower and flax, 18:1-CoA and 18:2-CoA are the preferred acyl donors and are incorporated into PC at similar rates (Stymne et al., 1981, 1984, 1985). Linolenate (α-18:3) is incorporated into PC at lower rates (Griffiths et al., 1985; Stymne et al, 1981, 1985), and saturated acyl groups are not efficiently utilized (Ichihara et al., 1995). Two recently cloned *Arabidopsis* LPCATs were shown to discriminate against 18:2 and 22:0, but utilized both saturated and monounsaturated C16 and C18 acyl groups (Shen et al., 2008). Similarly, LPCAT from microspore-derived cell suspension cultures of *B.napus* L. cv Jet Neuf could utilize 18:1, 18:0, and 16:0-CoA equally at concentrations around 20μM, but there was a preference for 18:1-CoA at higher

concentrations (Furukawa-Stoffer et al., 2003). Thus, the *in vitro* assay conditions should be considered when interpreting and comparing LPCAT specificity studies.

The reverse reaction of LPCAT is believed to be one mechanism by which acyl groups can be returned from to the acyl-CoA pool from PC (Stymne et al., 1984). *In vitro* studies with safflower microsomes have shown that in the presence of CoA and an acyl-CoA binding protein (such as BSA), the radiolabel from *sn*-2[^{14}C]18:2-PC was transferred to the acyl-CoA pool (Stymne et al., 1984). Subsequent radioisotope feeding studies have suggested that acyl exchange occurs at both the *sn*-1 and *sn*-2 positions (Bates et al., 2007), although it is unclear whether the LPCAT acts at both positions and how much acyl editing it accounts for. The reverse reaction of LPCAT is less efficient than the forward reaction and corresponds to about 5% of the acylation rate (Ichihara et al., 1995). No data are currently available on substrate specificity of the LPCAT reverse reaction in plants, but experiments in rat liver microsomes have demonstrated that 20:4, 18:0, and 18:2-CoA were preferentially synthesized from endogenous lipids under ATP-independent conditions (Sugiura et al., 1995). Thus, it is reasonable to expect that substrate specificities of both the forward and reverse reactions of LPCAT likely play a role in acyl-editing, and a more comprehensive understanding of these preferences may be key to overcoming substrate dichotomy in transgenic plants. The involvement of acyl-CoA binding proteins in mediating the reverse reaction of LPCAT may also indirectly

influence the substrate specificities of the reaction, since the binding proteins may have varying affinities for different acyl-CoA species, which would influence the substrate availability for the LPCAT reaction.

PDAT substrate specificity has only been characterized in a few species but it appears to exhibit some preference for polyunsaturated or unusual acyl groups. Ricinoleoyl, vernoloyl, 18:3 and 20:4 groups at the *sn*-2 position of PC were more efficiently utilized by PDAT in *Arabidopsis* microsomes compared to 18:2, 18:1 and 10:0, while 18:0 and 22:1 were excluded from the reaction (Stahl et al., 2004). It appears that PDAT plays a role in membrane maintenance by channelling unusual acyl groups from phospholipids into storage lipids. In some plants that produce unusual fatty acids (*Ricinus communis, Crepis palaestina*), PDAT has been shown to play a major role in removing these unusual acyl groups from membrane phospholipids, while others (*Crepis rubra, Euphorbia lagascae*) rely more on different mechanisms (Banas et al., 2000). In *Arabidopsis,* it was demonstrated that PDAT is not likely a major determinant of TAG composition or content in developing seeds (Mhaske et al., 2005). It is possible that other acyl-editing mechanisms, such as the Land`s Cycle (Lands et al., 1960), or exchange of backbones between PC and diacylglycerol catalyzed by choline phosphotransferase contribute to trafficking of acyl groups between PC and TAG in these species. The

relative contribution of various acyl-editing mechanisms to overall TAG composition, however, has yet to be completely dissected.

Future Outlook

Recent attempts to produce unusual fatty acids in transgenic plants have faced a number of challenges, which have the potential to drive major advances in our understanding of plant lipid metabolism. Earlier studies on the substrate preferences of various acyltransferases have often been limited to the use of crude microsomal preparations. Today, the ability to study acyltransferases in recombinant systems has revealed differences in substrate utilization between previously undistinguishable acyltransferase isoforms and is already providing the necessary insight to increase the levels of unusual fatty acids obtainable in transgenic plants. Similarly, an emerging understanding of the role of acyl-editing mechanisms and their associated specificities will be critical to overcoming existing bottlenecks in acyl group trafficking between various substrate pools. Insights into the structure-function and evolutionary relationships between various acyltransferases may also reveal new targets for biotechnological modification of plant lipids.

Acknowledgements

The authors acknowledge the support (to RJW) of the Natural Sciences and Engineering Research Council of Canada, the Alberta Agricultural Research Institute, AVAC, Ltd., the Alberta Crop Industry Development Fund and the Canada Research Chairs Program in relation to his research on acyltransferase action. We thank Dr. Disa Brownfield for her assistance in proofreading the manuscript.

Tables and Figures

Table 1. Biotechnological approaches to modifying seed oil content and composition through the use of acyltransferases in transgenic plants

Host species	Source species	Upregulated Gene	Phenotype	Reference
Brassica napus, Tropaeolum majus	*B. napus, Arabidopsis thaliana*	*DGAT1*	Increased seed oil content	(1, 2, 3)
A. thaliana, N. tobaccum	*A. thaliana*	*DGAT1*	Increased seed oil content, seed weight	(4, 5)
Glycine max	*Umbelopsis ramanniana*	*DGAT2*	Increased seed oil content	(6)
A. thaliana	*Ricinus communis*	*FAH12* *DGAT2*	Up to 30% hydroxyl fatty acids	(7)
B. napus	*Cocos nucifera*	*LPAAT* *12:0 ACP thioesterase*[a]	Increased accumulation of 12:0 at the *sn*-2 position from 5-30%.	(8)
B. napus	*Limnanthes spp., yeast*	*LPAAT*	Increased incorporation of 22:1 in *sn*-2 position	(9-12)
A. thaliana	*E. coli; C. tinctorius*[b]	*GPAT*	Increased seed oil content	(13)

[a]*Umbellularia californica* ACP-thioesterase

[b]*C. tinctorius* (safflower) plastidial GPAT

References:

1- Taylor et al. (2009)
2- Weselake et al. (2008)
3- Xu et al. (2008)
4- Bouvier-Nave et al. (2000)
5- Jako et al. (2001)
6- Lardizabal et al. (2008)
7- Burgal et al. (2008)
8- Knutzon et al. (1999)
9- Brough et al. (1996)
10- Lassner et al. (1995)
11- Taylor et al. (2002)
12- Zou et al. (1997)
13- Jain et al. (2000)

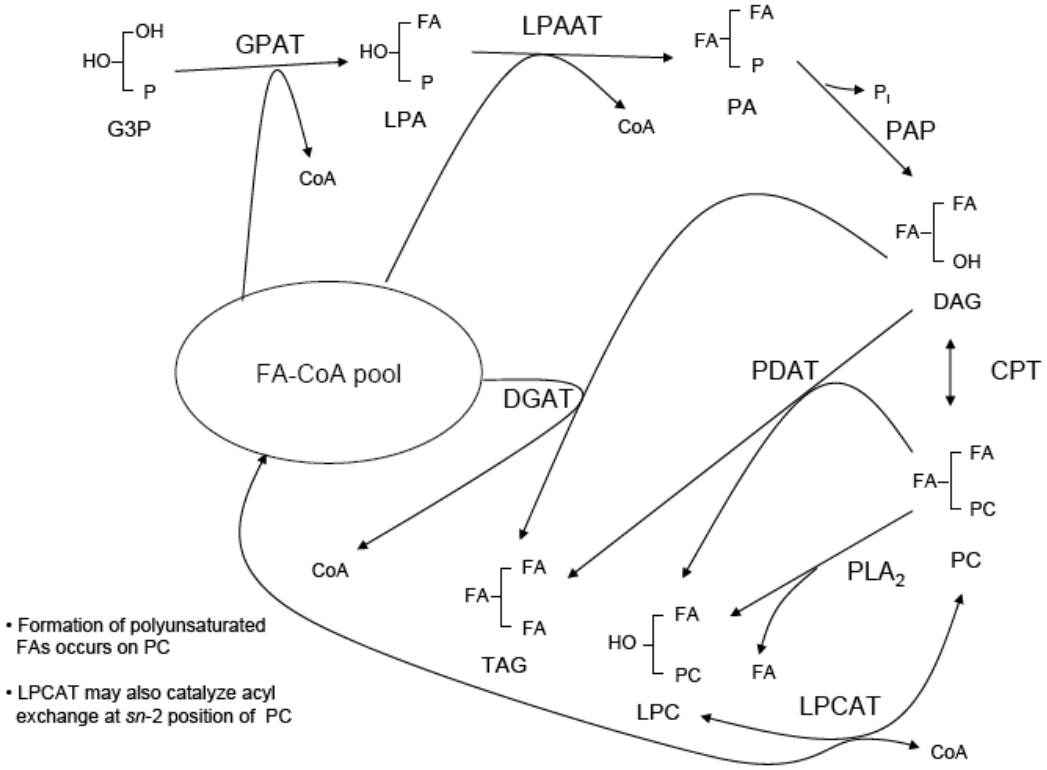


Figure 1. Generalized scheme for triacylglycerol (TAG) assembly in developing seeds of oleaginous plants. Abbreviations: CoA, coenzyme A; CPT, cholinephosphotransferase; DAG, *sn*-1, 2-diacylglycerol; FA, fatty acid; FA-CoA, fatty acyl-coenzyme A; G3P, *sn*-glycerol-3-phosphate; GPAT, *sn*-glycerol-3-phosphate acyltransferase; LPA, lysophosphatidic acid; LPAAT, lysophosphatidic acid acyltransferase; LPC, lysophosphatidylcholine; LPCAT, lysophosphatidylcholine acyltransferase; PA, phosphatidic acid; PAP, phosphatidic acid phosphatase; PC, phosphatidylcholine; PDAT, phospholipid:diacylglycerol acyltransferase; PLA_2, phospholipase A_2. Based on (Stymne et al., 1987) and (Weselake, 2005). PDAT may also catalyze the the synthesis of TAG and monoacylglycerol from two molecules of DAG (Ghosal et al., 2007).

References

Bafor M, *et al.*(1990) Properties of the glycerol acylating enzymes in microsomal preparations from the developing seeds of safflower (*Carthamus tinctorius*) and turnip rape (*Brassica campestris*) and their ability to assemble cocoa-butter type fats. *Journal of the American Oil Chemists Society*;67(4):217-25.

Banas A, *et al.*(2000) The involvement of phospholipid:diacylglycerol acyltransferases in triacylglycerol production. *Biochemical Society Transactions*;28:703-5.

Bates PD, *et al.*(2007) Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. *J Biol Chem*;282(43):31206-16.

Beisson F, *et al.*(2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *Plant Cell*;19(1):351-68.

Bernerth R, *et al.*(1990) Utilization of erucoyl-CoA by acyltransferases from developing seeds of *Brassica napus* L. involved in triacylglycerol biosynthesis. *Plant Science*;67(1):21-8.

Bouvier-Nave P, *et al.*(2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur J Biochem;*267(1):85-96.

Brokerhoff H, *et al.*(1966) Stereospecific analyses of several vegetable fats. *Journal of Lipid Research*;7:62-4.

Brown AP, *et al.*(2002) Substrate selectivity of plant and microbial lysophosphatidic acid acyltransferases. *Phytochemistry*;61(5):493-501.

Brough CL, *et al.*(1996) Towards the genetic engineering of triacylglycerols of defined fatty acid composition: Major changes in erucic acid content at the *sn*-2 position affected by the introduction of a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase from *Limnanthes douglasii* into oil seed rape. *Molecular Breeding*;2(2):133-42.

Burgal J, *et al.*(2008) Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnol J*;6(8):819-31.

Cahoon EB, *et al.* (2007) Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr Opin Plant Biol*;10(3):236-44.

Cao YZ, *et al.*(1990) Lysophosphatidate acyltransferase in the microsomes from maturing seeds of meadowfoam (*Limnanthes alba*). *Plant Physiology*;94(3):1199-206.

Christie WW, *et al.*(1991) Stereospecific analysis of triacyl-*sn*-glycerols via resolution of diastereomeric diacylglycerol derivatives by high performance liquid chromatography on silica. *Journal of the American Oil Chemists Society*;68(10):695-701.

Dahlqvist A, *et al.*(2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proceedings of the National Academy of Sciences of the United States of America*;97(12):6487-92.

Davies HM, *et al.*(1995) Lysophosphatidic acid acyltransferase from immature coconut endosperm having medium chain length substrate specificity. *Phytochemistry*;39(5):989-96.

Drexler H, *et al.*(2003) Metabolic engineering of fatty acids for breeding of new oilseed crops: strategies, problems and first results. *J Plant Physiol*;160(7):779-802.

Dutta PC, *et al.*(1992) Utilization of petroselinate (C18:1□6) by glycerol acylation enzymes in microsomal preparations of developing embryos of carrot (*Daucus carota* L.), safflower (*Carthamus tinctorius* L.) and oil rape (*Brassica napus* L.). *Plant Science*;81(1):57-64.

Frentzen M.(1993) Acyltransferases and triacylglycerols. In: *Lipid Metabolism in plants*.(Moore Jr., T.S., editors)p. 195-230 CRC Press.

Furukawa-Stoffer TL, *et al.*(2003) Properties of lysophosphatidylcholine acyltransferase from *Brassica napus* cultures. *Lipids*;38(6):651-6.

Ghosal A, *et al.*(2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyl transferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate specificities. Biochimica et Biophysica Acta-Molecular and Cell Biology of Lipids;1771(12):1457-63.

Griffiths G, *et al.*(1985) The acylation of *sn*-glycerol-3-phosphate and the metabolism of phosphatidate in microsomal preparations from the developing cotyledons of safflower (*Carthamus tinctorius* L.) seed. *Biochemical Journal*;230(2):379-88.

Harwood JL.(2005) Fatty acid biosynthesis. In: *Plant Lipids: Biology, Utilization and Manipulation*.(Murphy DJ, editor)p. 27-66 Blackwell Publishing.

He XH, *et al.*(2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. *Lipids*;39(4):311-8.

Ichihara K, *et al.*(1982) Lipid metabolism in safflower. 4. Some properties of diacylglycerol acyltransferase in a particulate fraction from maturing safflower seeds. *Phytochemistry*;21(8):1895-901.

Ichihara K, *et al.*(1988) Diacylglycerol acyltransferase in maturing safflower seeds - its influences on the fatty acid composition of triacylglycerol and on the rate of triacylglycerol synthesis. *Biochimica et Biophysica Acta*;958(1):125-9.

Ichihara K, *et al.*(1995) 1-Acylglycerophosphocholine O-acyltransferase in maturing safflower seeds. *Planta*;196(3):551-7.

Jako C, *et al.*(2001) Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol*;126(2):861-74.

Jain RK, *et al.*(2000) Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. *Biochemical Society Transactions*;28:958-61.

Katavic V, *et al.*(2001) Improving erucic acid content in rapeseed through biotechnology: What can the *Arabidopsis* FAE1 and the yeast SLC1-1 genes contribute? *Crop Science*;41(3):739-47.

Katavic V, *et al.*(1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. *Plant Physiology*;108(1):399-409.

Knutzon DS, *et al.*(1999) Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the sn-2 position of triacylglycerols in lauric rapeseed oil and can increase total laurate levels. *Plant Physiology*;120(3):739-46.

Lands WEM.(1960) Metabolism of glycerolipids .2. enzymatic acylation of lysolecithin. *Journal of Biological Chemistry*;235(8):2233-7.

Lardizabal K, *et al.*(2008) Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiol*;148(1):89-96.

Larson TR, *et al.*(2002) Acyl-CoA profiles of transgenic plants that accumulate medium-chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. *Plant J*;32(4):519-27.

Lassner MW, *et al.*(1995) Lysophosphatidic acid acyltransferase from meadowfoam mediates insertion of erucic acid at the *sn*-2 position of triacylglycerol in transgenic rapeseed oil. *Plant Physiology*;109(4):1389-94.

Laurent P, *et al.*(1992) Organ- and development-specific acyl coenzyme A lysophosphatidate acyltransferases in palm and meadowfoam. *Plant Physiol*;99(4):1711-5.

Li Y, *et al.*(2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc Natl Acad Sci U S A*;104(46):18339-44.

Lisa M, *et al.*(2008) Triacylglycerols profiling in plant oils important in food industry, dietetics and cosmetics using high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry. *J Chromatogr A*;1198-1199:115-30.

Lohden I, *et al.*(1992) Triacylglycerol biosynthesis in developing seeds of *Tropaeolum majus* L. and *Limnanthes douglasii* R. Br. *Planta*;188(2):215-24.

Lung SC, *et al.*(2006) Diacylglycerol acyltransferase: a key mediator of plant triacylglycerol synthesis. *Lipids*;41(12):1073-88.

Martin BA, *et al.*(1983) Properties of diacylglycerol acyltransferase from spinach leaves. *Lipids;*18(1):1-6.

Mhaske V, *et al.*(2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid:diacylglycerol transacylase gene (At5g13640). *Plant Physiol Biochem*;43(4):413-7.

Oo KC, *et al.*(1989) Lysophosphatidate acyltransferase activities in the microsomes from palm endosperm, maize scutellum, and rapeseed cotyledon of maturing seeds. *Plant Physiol*;91(4):1288-95.

Perry HJ, *et al.*(1993) Changes in the lipid content of developing seeds of *Brassica napus*. *Phytochemistry*;32(6):1411-5.

Princen LH, *et al.*(1984) Development of new crops for industrial raw materials. *Journal of the American Oil Chemists Society*;61(2):281-9.

Shen W, *et al.* Functional study of lysophosphatidylcholine acyltransferase in *Arabidopsis* (poster presented at 18th International Symposium on Plant Lipids, Bordeaux, France, July 20-25, 2008.

Shockey JM, *et al.*(2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell*;18(9):2294-313.

Siloto RM, *et al.*(2009) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: development and characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiol Biochem*;47(6):456-61。

Sorensen BM, *et al.*(2005) Storage lipid accumulation and acyltransferase action in developing flaxseed. *Lipids*;40(10):1043-9.

Stahl U, *et al.*(2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from *Arabidopsis*. *Plant Physiology*;135(3):1324-35.

Stymne S, *et al.*(1987) Triacylglycerol biosynthesis. In: *The Biochemistry of Plants vol. 9, lipids: structure and function*.(Stumph PK, editor)p. 175-214 Academic Press.

Stymne S, *et al.*(1981) Acyl exchange between oleoyl-CoA and phosphatidylcholine in microsomes of developing soya bean cotyledons and its role in fatty acid desaturation. *Lipids*;16(5):298-305.

Stymne S, *et al.*(1985) Oil synthesis *in vitro* in microsomal membranes from developing cotyledons of *Linum usitatissimum* L. Planta;164(1):101-4.

Stymne S, *et al.*(1984) Evidence for the reversibility of the acyl-CoA-Lysophosphatidylcholine Acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius L.*) cotyledons and rat liver. *Biochemical Journal*;223(2):305-14.

Sugiura T, *et al.*(1995) Coenzyme A dependent modification of fatty acyl chains of rat liver membrane phospholipids - possible involvement of ATP independent acyl-CoA synthesis. *Journal of Lipid Research*;36(3):440-50.

Suh MC, *et al.*(2002) What limits production of unusual monoenoic fatty acids in transgenic plants? *Planta*;215(4):584-95.

Sun C, *et al.*(1988) Acyl Coenzyme A preference of the glycerol phosphate pathway in the microsomes from the maturing seeds of palm, maize, and rapeseed. P*lant Physiol*;88(1):56-60.

Taylor DC, Zhang Y, Kumar A, Francis T, Giblin ME, Barton DL, et al. Molecular modification of triacylglycerol accumulation by over-expression of *DGAT1* to produce canola with increased seed oil content under field conditions. *Botany*. In press 2009.

Taylor DC, *et al.*(2002) Field testing of transgenic rapeseed cv. Hero transformed with a yeast *sn*-2 acyltransferase results in increased oil content, erucic acid content and seed yield. *Molecular Breeding*;8(4):317-22.

Taylor DC, *et al.*(1991) Triacylglycerol bioassembly in microspore-derived embryos of *Brassica napus* L. cv Reston. *Plant Physiology*;97(1):65-79.

Taylor DC, *et al.*(1992) Biosynthesis of acyl lipids containing very-long chain fatty acids in microspore-derived and zygotic embryos of *Brassica napus* L. cv Reston. *Plant Physiology*;99(4):1609-18.

Voelker T, *et al.*(2001) Variations in the biosynthesis of seed storage lipids. *Annu Rev Plant Physiol Plant Mol Biol*;52:335-61.

Voelker TA, *et al.*(1996) Genetic engineering of a quantitative trait: Metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. *Plant Journal*;9(2):229-41.

Vogel G, *et al.*(1996) Cholinephosphotransferase and diacylglycerol acyltransferase - Substrate specificities at a key branch point in seed lipid metabolism. *Plant Physiology*;110(3):923-31.

Weselake RJ.(2005) Storage lipids. In: *Plant Lipids: Biology, Utilization and Manipulation*.(Murphy DJ, editor)p. 162-221 Blackwell Publishing.

Weselake RJ, *et al.*(2008) Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. *Journal of Experimental Botany.*

Wiberg E, *et al.*(2000) The distribution of caprylate, caprate and laurate in lipids from developing and mature seeds of transgenic *Brassica napus* L. *Planta*;212(1):33-40.

Williams JP, *et al.*(2000) The role of phosphatidylcholine in fatty acid exchange and desaturation in *Brassica napus* L. leaves. *Biochemical Journal*;349:127-33.

Xu J, *et al.*(2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. *Plant Biotechnology Journal*. 6 (8): 799-818.

Yu KS, *et al.*(2006) Diacylglycerol acyltransferases from *Vernonia* and *Stokesia* prefer substrates with vernolic acid. *Lipids*;41(6):557-66.

Zou J, *et al.*(1997) Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast *sn*-2 acyltransferase gene. *Plant Cell*;9(6):909-23.

Zheng P, *et al.*(2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nature Genetics*;40(3):367-72.

Zheng Z, *et al.*(2003) *Arabidopsis* AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell*;15(8):1872-87.